Universities, Spatial Misallocation, and Growth

Santiago Caicedo* Xiao Ma† Yiran Zhang‡

March 2026

Abstract

We study how spatial spillovers from universities shape aggregate economic growth. Using Chinese data on manufacturing firms, patents, and land transactions, we show that proximity to universities raises innovation efficiency, primarily through knowledge diffusion and access to university-trained talent. New firms benefit disproportionately from these spillovers yet are less likely to locate nearby, where incumbents occupy scarce land. We build a quality-ladder model in which firms choose both innovation effort and location. Creative destruction reallocates product lines but not physical space: an innovating entrant inherits the incumbent's product but not its land, generating spatial misallocation absent from standard growth models. The social planner concentrates more R&D effort near universities than the market does, and this spatial distortion compounds the familiar under-investment in innovation. Calibrating to Chinese data, preliminary counterfactual exercises suggest that university spillovers are a sizable source of growth. Eliminating the proximity advantage in innovation productivity would cut the balanced growth rate by more than half, while land reallocation policies yield additional but more modest gains.

JEL Codes: O31, O41, R30

Key words: Knowledge Spillovers, Spatial Misallocation, Innovation Geography

*Caicedo: Northeastern University, sa.caicedo@northeastern.edu.

†Ma: Peking University HSBC Business School, xiaoma@phbs.pku.edu.cn; xiaomaecon@gmail.com.

‡Zhang: Fudan University, yiran_zhang@fudan.edu.cn; yiranzhangecon@gmail.com

1 Introduction

Disseminating knowledge is widely regarded as a fundamental role of universities. Many empirical studies find that proximity to universities boosts firms' ability to develop new knowledge, significantly enhancing nearby firms' innovation activities (e.g., Jaffe 1989, Andrews 2017, Kantor & Whalley 2014, Hausman 2020). This is also evident in real-world examples such as Stanford, which has evolved into a major hub of innovation, attracting entrepreneurs and companies to the surrounding area known as "Silicon Valley." However, whereas knowledge spillovers have long been considered a crucial driver of aggregate economic growth in the macroeconomic literature (e.g., Romer 1990, Jones 1995, Lucas & Moll 2014), there is a lack of quantitative analysis examining the relationship between knowledge spillovers from universities and their impact on aggregate economic growth.

In this paper, we address this gap by developing a quantitative growth framework that incorporates creative destruction and a novel spatial dimension—specifically, proximity to universities, firm location choices, and land availability. Using detailed Chinese data on manufacturing firms, patents, and geography, we provide new evidence on how university proximity shapes innovation through both direct channels (co-invention and patent transfers) and indirect spillovers (talent pools and knowledge diffusion). We then build a quality-ladder model in which firms choose not only how much to innovate but also where to locate, subject to land availability constraints. The model reveals a form of spatial misallocation that standard growth models overlook: when an entrant innovates and displaces an incumbent, it takes over the incumbent's product line but not the incumbent's physical location. This decoupling between product-space competition and physical-space allocation means that more productive innovators are systematically underrepresented near universities, where innovation spillovers are strongest.

We begin by documenting the empirical relationship between firms' innovation and their proximity to universities. China serves as a relevant context due to its rapid increase in innovation—domestic invention applications surged over 15-fold from 2000 to 2015, with universities and research institutes contributing 30%—highlighting the significance of knowledge diffusion (Wei, Xie & Zhang 2017). It also features rich spatial variation, encompassing many regions and universities.

Our empirical analysis uncovers several key patterns. First, within cities,[1] firms located closer to universities benefit more from direct interactions with academic institutions: they engage in more co-invention with universities and receive more patent transfers and licenses. Beyond these direct channels, proximity also raises firms' own innovation productivity. Firms nearer to universities produce more patents per R&D worker, though patents per R&D expenditure remain relatively constant across distances. We investigate two channels behind this indirect spillover—access to a better talent pool and knowledge diffusion from university research—and find that both contribute significantly, with knowledge diffusion accounting for the larger share. Using instrumental variables that exploit variation from geographically distant universities, we show that these results are robust to endogeneity concerns. After controlling for both channels, the innovation premium associated with proximity to universities nearly vanishes, suggesting that talent and knowledge diffusion together account for most of the observed proximity advantage.

Second, we examine how these patterns differ for new versus incumbent firms. New firms tend to innovate more efficiently than incumbents, especially when located farther from universities, suggesting they make more effective use of their R&D workforce. Moreover, new firms benefit disproportionately from knowledge diffusion, indicating they are particularly responsive to academic spillovers. Yet despite this responsiveness, new firms are less likely to locate near universities, and firm relocations are rare, underscoring the stickiness of initial location choices.

Third, using employer-employee data from Revelio Labs, we document that college graduates tend to remain near the universities from which they graduated. Graduates are roughly 3.5 times more likely to work in their college city than a random assignment would predict, and even conditional on staying in the same city, local graduates work closer to universities than non-local graduates. This sorting of talent reinforces the innovation advantage of locations near universities and provides an additional mechanism through which university proximity raises firms' research productivity.

These empirical findings point to a potential spatial misallocation of innovation. New firms, despite being more efficient innovators, are less likely to be near universities, where proximity

[1]We focus on within-city proximity to universities due to the localized nature of knowledge spillovers documented in the literature (Jaffe, Trajtenberg & Henderson 1993, Murata, Nakajima, Okamoto & Tamura 2014, Kerr & Kominers 2015, Berkes, Gaetani & Mestieri 2025).

boosts innovation through better access to talent and knowledge. To formalize and measure this misallocation, we develop an endogenous growth model grounded in the classical quality-ladders framework (Aghion & Howitt 1992, Grossman & Helpman 1991). In the model, a fixed measure of product lines are operated by incumbent firms, and new entrants engage in innovation to improve product quality and displace incumbents. Entrants can choose their proximity to universities, which influences their innovation efficiency due to varying exposure to academic knowledge. While all firms have an incentive to locate near universities to benefit from higher spillover effects, limited land availability constrains this behavior: increased demand for proximity drives up land prices, reducing firms' profits and limiting clustering.

To isolate the model's key mechanism, we first analyze a simplified version with two locations that differ in innovation productivity and land endowments. The model highlights three forces: knowledge spillovers from universities raise innovation productivity in nearby locations; limited land near universities constrains how many firms can locate there; and creative destruction determines land availability, but freed land does not follow the innovator. Comparing the competitive equilibrium to the social planner's solution yields three results. First, the market equalizes R&D effort across locations, even though the social return to R&D is higher near the university; the planner optimally concentrates more R&D labor in the productive location. Second, the planner sends a larger share of entrants to the near-university location, accepting greater land crowding for faster growth. Third, these two margins compound, so that the planner's innovation gap between locations strictly exceeds the market's. The spatial dimension of inefficiency goes beyond the standard under-investment result in Schumpeterian models: not only is overall innovation below the social optimum, but both its geographic composition and its intensity composition are distorted.

We then extend the model to a fully quantitative framework with multiple regions and distance bins. The quantitative model incorporates innovation by both entrants and incumbents, an innovation lifecycle in which firms' R&D efficiency can decline over time, and separate markets for skilled and unskilled labor. We calibrate the model using Chinese firm-level and land data, combining external estimates with internally calibrated parameters via the simulated method of moments (SMM). Key targeted moments include TFP growth, R&D intensity by firm age, patenting across distance bins, and land prices. Our estimates align well with values found in the literature, and the model captures the main documented patterns, including higher innovation efficiency near universities and greater R&D efficiency

in new firms, matching both targeted and non-targeted data moments.

Our counterfactual exercises are still work in progress, but preliminary results illustrate the quantitative importance of university spillovers and land constraints. Reducing innovation productivity in all locations to the level of the farthest bin lowers the balanced growth rate from 3% to about 1%, with a welfare decline of over 6%. Reallocating land toward locations closer to universities initially raises growth and welfare, though the gains diminish as increased competition for land drives up prices. We are also exploring additional policies, including industrial land allocation and university funding, to further examine their impacts on spatial misallocation and regional disparities.

Literature. This paper relates to several strands of the literature. First, this paper connects with the macro literature on the role of education in driving economic growth. Seminal work by Romer (1990) and Aghion & Howitt (1992) models human capital as a crucial input to the research sector, which generates new ideas and fosters technological progress. A substantial macro literature has since documented the positive cross-country relationship between higher education and future technology growth (e.g., Barro 1991, Benhabib & Spiegel 1994, Furman, Porter & Stern 2002, Valero & Reenen 2019). Universities are a vital component of the education system, playing a key role in building talent and disseminating knowledge. While recent studies have highlighted how the supply of talent from universities contributes to economic growth (Akcigit, Pearce & Prato 2020, Ma 2024), there remains limited understanding of how knowledge spillovers from universities affect aggregate growth—an issue this paper seeks to address.

Our paper also contributes to a broad literature examining the impact of university spillovers, with a particular emphasis on identifying causal relationships. For instance, Andrews (2017) employs a "finalist" empirical strategy that leverages historical site selection decisions of U.S. universities to estimate causal effects. Similarly, Aghion, Bergeaud, Lequien & Melitz (2017) uses political shifts in education funding to assess the impact of colleges on innovation, and Hausman (2020) exploits the enactment of the Bayh-Dole Act as a natural experiment. Although these studies successfully exploit quasi-experimental variation to estimate the regional effects of universities, they largely focus on reduced-form relationships, leaving the aggregate implications for growth and policy insufficiently explored. This paper fills that

gap by integrating empirical evidence with a general equilibrium spatial growth model to quantitatively evaluate the role of university spillovers in driving aggregate economic growth and to assess policies that can amplify their impact.

Finally, our paper contributes to the quantitative literature on the geography of innovation and knowledge diffusion (e.g., Eaton & Kortum 2001, Alvarez, Buera & Lucas 2013, Arkolakis, Ramondo, Rodrıguez-Clare & Yeaple 2018, Desmet, Nagy & Rossi-Hansberg 2018, Buera & Oberfield 2020, Cai, Caliendo, Parro & Xiang 2022, Liu & Ma 2023), as surveyed by Lind & Ramondo (2019). While much of this work treats knowledge spillovers as exogenous or emphasizes their transmission through trade and migration networks, we take a different approach. Building on this literature, we develop a quantitative spatial growth model that highlights the role of universities in generating spatial spillovers depending on firms' proximity, and we examine how these spillovers influence both the spatial distribution of economic activity and the aggregate growth rate.

The rest of the paper is organized as follows. In Section 2, we describe our data and examine the empirical relationship between innovation and proximity to universities. Guided by these findings, we develop the Schumpeterian growth model incorporating land and firm location choices in Section 3. Section 4 presents our counterfactual exercises along with preliminary results. Finally, we conclude in Section 5.

2 Innovation and Distance to Universities

In this section, we document key empirical patterns linking firms' innovation activity to their proximity to universities. We examine how this relationship differs between new entrants and incumbent firms, and how it relates to their location decisions. These empirical facts serve as motivation for our theoretical model and provide discipline for our quantitative framework.

2.1 Data Sources

We begin by briefly describing the datasets of innovation, firm outcomes, and geographic information employed in our analysis.

Patent Data. We use Chinese patent application data assembled by the China National Intellectual Property Administration. It provides a comprehensive source of data that can be used to gain insights into the innovation activities of Chinese firms, covering detailed information on each patent, including the unique application number, application date, grant date, inventors, and International Patent Classification (IPC) code. We focus on all invention applications[2] that were applied for and eventually granted during the 1985–2021 period. The dataset covers patents filed by firms, universities, and research institutes, allowing us to identify co-inventions between firms and universities, patent transfers and licenses from universities to firms, and to construct measures of technological similarity between firms and universities based on their respective IPC patent portfolios.

Firm-level Data. For firm-level data, we integrate multiple sources. The primary dataset is China's Annual Survey of Manufacturing (ASM), an administrative dataset covering firm demographics and balance sheets. The ASM surveys 200,000 to 400,000 firms per year, representing over 90% of China's manufacturing output. We use data from 2000 to 2013, which includes firm names, value added, sales, employment, capital stock, 4-digit industry affiliation, and registration type for firms above a sales threshold.[3]

One limitation of the ASM data is the lack of R&D data,[4] particularly lacking detailed information on R&D personnel. To complement the ASM data, we use the Survey of R&D Firms from 2008 to 2013, which covers industrial firms engaged in R&D and includes roughly 290 thousand firm-year observations. Using firm names, we match these firms to those in the ASM, achieving a linkage rate of over 95% of firm-year observations in the R&D survey. This dataset offers rich details on R&D activities, including workforce composition, educational background, funding sources, and expenditure components.

We link the patent data with firm-level data using firms' names, after cleaning and standardizing firm names. The specifics regarding the datasets and the process of consolidating

[2]China recognizes three types of patentsinvention, utility model, and design—with invention patents generally considered the most technology-intensive among them.

[3]From 2000 to 2007, the sales threshold was 5 million RMB (approximately $600,000), and the sample included all state-owned firms. After 2011, the threshold increased to 20 million RMB for all firms. Since the dataset covers all medium-sized and large firms, it provides a reliable measure of aggregate manufacturing sales by industry. Brandt, Van Biesebroeck & Zhang (2012) estimates that firms below this threshold accounted for only 9.9% of total industrial output in 2004.

[4]In the ASM, data on R&D expenditures are available only for the periods 2001–2002 and 2005–2007.

firm names are outlined in Appendix A.1.

We enrich our dataset with geographic information by geocoding firm addresses using the GaoDe Maps API, obtaining longitude and latitude coordinates for each firm-year observation. For the 2008–2013 R&D survey subsample, this process yields a match rate of 96%. We also collect geographic data on universities from their official websites and on research institutes (we exclude research institutes affiliated with firms, focusing instead on those associated with universities, the military, or government agencies) identified in the patent dataset. We apply the geographic coordinates to calculate each firm's shortest distance to the nearest university or research institution.

Land Data. Finally, to capture the interaction between land availability and proximity to universities in our quantitative analysis, we compile data on plot-level land transactions in China from 2000 to 2023.[5] Each transaction record includes the city, geographic coordinates, land price, land area, and designated land use (e.g., residential, commercial, or industrial). We use the geographic coordinates to calculate each land's shortest distance to the nearest university or research institution.

Employer–Employee Data. To examine the location choices of university graduates, we use individual-level workforce data from Revelio Labs.[6] The dataset contains detailed work histories, including job titles, employer information, start and end dates, and educational background (institution name and graduation date) for individuals across a wide range of industries. We geocode company and university addresses and retain only firms and universities with available location information for individuals who received their college education in China. Each observation in the dataset corresponds to a job held by an individual at a specific firm. We restrict the sample to jobs with start dates no earlier than the individual's graduation date, yielding information on roughly 2.9 million jobs. Although the data may be subject to selection issues inherent to online professional profiles, it represents the largest available employer–employee dataset for China and allows us to study the spatial sorting of

[5]The data are sourced from `https://www.landchina.com/`, a government platform that records and publishes land transaction information.

[6]Revelio Labs collects and standardizes publicly available professional profiles to construct a comprehensive workforce database. The data are accessible through the Wharton Research Data Services (WRDS) platform. See `https://www.reveliolabs.com/` for details.

college graduates relative to their alma maters.

2.2 Firms' Innovation Gains from Proximity to the University

We first use our compiled data to document how firms' innovation outcomes differ based on their proximity to universities. We begin with direct benefits from firm–university interactions, then turn to indirect spillover effects, and finally examine how these patterns differ for new versus incumbent firms.

Direct gains through co-innovation and transfers. Firms can benefit from universities through direct interactions, such as co-invention, where both universities and firms are listed as patent assignees, or through the transfer of patents, in which universities convey patent applications or usage rights to firms, as well as through the licensing of patent usage rights. In Table 1, we report counts of firm self-innovation patents, co-invented patents, patents transferred from universities to firms, and university patents that were neither transferred nor licensed to firms. These counts are provided for both the full patent sample and the subset corresponding to our firm dataset, with brackets indicating each category's share of all patents held by firms. The results suggest that co-invention, patent transfers, and licensing are economically meaningful, representing 7.7% of all patents held by firms in the full dataset and 6.2% in the subset matched to our R&D firm sample for 2008–2013.

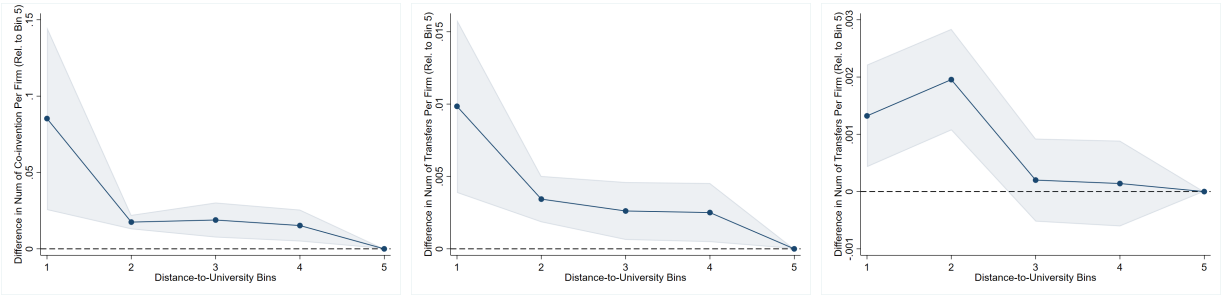
In principle, firms located closer to universities may benefit more from direct interaction with academic institutions. To examine this, within each city-year we sort firms into five equally sized bins, ($r = 1, ..., 5$), based on their distance to the nearest university (or research institute). Firms in the most distant bin serve as the benchmark group. Figure 1 shows that firms located closer to universities gain significantly more from co-invention, patent transfers, and licensing. The figure also indicates that these proximity-related benefits are generally stronger for co-invention than for patent transfers or licenses.

Table 1: Number of Invention Patents

			University		
Application year	Firm self-invention	Co-invention (firm&univ.)	Transferred	Licensed	Other
1985–2021	2,011,398 (92.3%)	84,123 (3.9%)	70,276 (3.2%)	13,219 (0.6%)	960,636
2008–2013	560,549 (90.3%)	26,923 (4.3%)	28,096 (4.5%)	5,372 (0.9%)	241,643
Our firm sample 2008–2013:					
2008–2013	296,457 (93.9%)	12,889 (4.1%)	4,258 (1.3%)	2,075 (0.7%)	–

Notes: Counts of granted invention patents by type of firm–university interaction. "Co-invention" refers to patents listing both a firm and a university as assignees. "Transferred" and "Licensed" refer to patents whose application or usage rights were conveyed from universities to firms. Percentages in brackets indicate each category's share of all patents for which firms held user rights during the same period. "Other" denotes university patents that were neither transferred nor licensed to firms.

Figure 1: Distance to University and Direct Benefits



(a) Co-innovation (b) Transfers from universities (c) Licenses from universities

Notes: Each panel plots the difference in the number of direct university interactions relative to firms in the farthest distance bin (baseline group). Within each city-year, firms are divided equally into five bins based on their minimum distance to any university or research institute. All regressions control for city and year fixed effects. Shaded areas represent 95% confidence intervals based on robust standard errors.

Indirect spillover effects. Proximity to universities provides not only benefits through direct firm–university interactions but also indirect spillovers by enhancing firms' own innovation. We detect such spillover effects by focusing on how the number of self-innovation inventions varies with distance to universities.

Firms can, in principle, benefit from universities through multiple channels. We focus on two fundamental functions of universities—teaching and research—and therefore examine how access to a better talent pool and knowledge diffusion improves nearby firms' research productivity. We capture the talent pool by the share of university graduates among a firm's R&D personnel, defined as $S_{i,t}$. To measure the strength of knowledge diffusion, we take advantage of the spillover measure suggested by Bloom, Schankerman & Van Reenen

(2013) and extend it by taking into account the geographic barriers to knowledge spillovers. The idea of this measure is that a firm potentially benefits more from other firms whose technology bundles are more similar to the firm's technology bundle. We define the vector $\mathbf{T}_i = (T_i^1, T_i^2, ..., T_i^{124})$, where T_i^τ is the share of firm i's patents in the technology class indexed by τ. We consider 124 3-digit IPC categories as technology classes. We similarly construct technology bundles for all universities. For firm i and university j, we then construct the similarity between their technology bundles according to Jaffe (1986):

$$\rho_{i,j} = \frac{\mathbf{T}_i \mathbf{T}_j'}{\left(\mathbf{T}_i \mathbf{T}_i'\right)^{1/2} \left(\mathbf{T}_j \mathbf{T}_j'\right)^{1/2}}. \tag{1}$$

Thus, the index $\rho_{i,j}$ ranges between 0 and 1 and is closer to 1 if firms i and university j have more patent applications in the same technology class. We compute $\rho_{i,j}$ between each local firm and each university, using their cumulative patent applications in our full patent data.

With this index, we can compute the technology spillover from universities for each firm i in year t,

$$\text{diffusion}_{i,t} = \sum_{j \in U} \rho_{i,j} \frac{patent_{j,t}}{d_{i,j}^{\eta}},$$

where $d_{i,j}$ is the distance between them, and $patent_{j,t}$ is the number of patent applications by university j in year t.[7] η governs the rate at which knowledge diffusion decays with distance. We set $\eta = 1$, implying that knowledge diffuses inversely with distance, consistent with citation patterns to university patents documented in Belenzon & Schankerman (2013).

We perform the following regression:

$$y_{i,t} = \beta_0 + \sum_{r=1}^{4} \beta_r bin_{r(i,t)} + \gamma S_{i,t} + \theta \log(1 + \text{diffusion}_{i,t}) + \phi \mathbf{X}_{i,t} + \delta_{c(i,t)} + \zeta_t + \epsilon_{i,t}. \tag{2}$$

$y_{i,t}$ is the outcome variable of interest. $bin_{r(i,t)}$ are the dummies of distance bins constructed before (we treat the farthest bin as the baseline group). As for some firms, $\text{diffusion}_{i,t}$ may contain zero values, we use $\log(1 + \text{diffusion}_{i,t})$ transformation. Appendix Table A.1 shows that our results remain robust to using the Inverse Hyperbolic Sine (IHS) transformation. The control variables $\mathbf{X}_{it}$ include firm characteristics, specially deciles of firm sales and

[7] Bloom et al. (2013) use the stock of R&D to measure the source of spillovers. As we lack R&D data for universities, we instead use the number of patent applications.

industry fixed effects to capture heterogeneity in innovation patterns across firm size and industries. $\delta_{c(i,t)}$ and ζ_t correspond to city and year fixed effects for firm i in time t.

Table 2 presents the regression results. In Column (1), we regress the number of self-innovation patents on the dummies of distance bins, without adding the two channels. It shows that firms located closer to universities file more invention applications per year. Because firms differ in the size of their R&D teams—which can in turn affect their innovation output due to scale effects—Column (2) examines a more informative measure, research productivity, defined as the number of self-innovation patents per 100 R&D employees. Consistent with Column (1), it shows that firms located closer to universities exhibit higher innovation productivity. In Column (3), we add the two channels—better talent pool and knowledge diffusion—and show that the share of university graduates and knowledge diffusion have a strong impact on firms' innovation productivity.

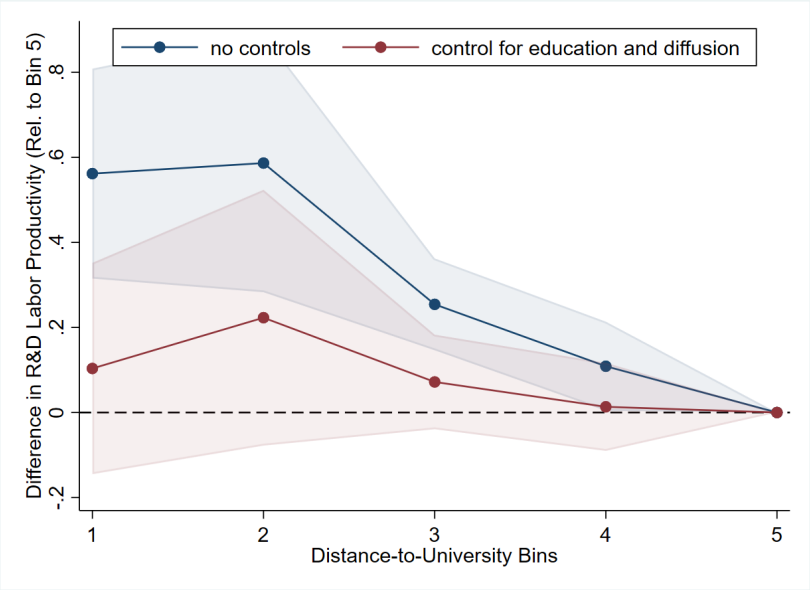
The share of university graduates and knowledge diffusion may be endogenous, as firms that are more efficient in research may deliberately locate closer to universities. To lessen the endogeneity concern, we construct two instruments by eliminating the direct impacts of nearby universities. Specifically, we construct two instruments as follows:

$$SI^{iv}_{i,t} = SI^{-i}_{s(i),t} \times \log\left(\sum_{j \in U \setminus \mathcal{C}_i} \frac{grad_{j,t}}{d^{\eta}_{i,j}}\right), \quad \text{diffusion}^{iv}_{i,t} = \sum_{j \in U \setminus \mathcal{C}_i} \rho_{i,j} \frac{patent_{j,t}}{d_{i,j}}. \tag{3}$$

$SI^{-i}_{s(i),t}$ represents the nationwide share of university graduates among R&D personnel in the firm's 2-digit industry, excluding the firm itself. It captures industry-specific reliance on high-skilled talent and is interacted with the firm's exposure to the university talent pool, measured as $\sum_{j \in U \setminus \mathcal{C}_i} \frac{grad_{j,t}}{d^{\eta}_{i,j}}$. Here, $\mathcal{C}_i$ denotes the set of nearby universities, which are excluded when constructing this exposure. Similarly, the instrument for knowledge diffusion is based on patents originating from more distant universities, $\sum_{j \in U \setminus \mathcal{C}_i} \rho_{i,j} \frac{patent_{j,t}}{d_{i,j}}$, again excluding nearby contributions. In the baseline regressions, $\mathcal{C}_i$ is defined as the set of universities located within 10 kilometers of firm i.

Column (4) shows the result of 2SLS regressions. We find that the share of university graduates and knowledge diffusion still have a strong impact on firms' innovation productivity. After controlling for these two channels, the impact of distance bins on innovation productivity nearly vanishes, as shown by Figure 2.

Figure 2: Productivity Advantage with Proximity to Universities



Notes: Estimated coefficients on distance-bin dummies from regressions of self-invention patents per 100 R&D workers. The blue series corresponds to Column (2) of Table 2 (no channel controls); the red series corresponds to Column (3) (controlling for the share of university graduates and knowledge diffusion). Distance bin 5 (farthest from universities) is the omitted category. Shaded areas represent 95% confidence intervals based on robust standard errors.

To assess the relative importance of the two channels, we proceed as follows. For each distance bin, we multiply the estimated coefficient on the share of university graduates by the difference in this share between that bin and the benchmark group (distance bin 5). We then divide this product by the decline in the coefficient on the distance-bin dummy relative to Column (2), which excludes channel controls. Using this approach, we find that improvements in the talent pool account for 9–15% of the combined effects of the two channels based on the OLS estimates in Column (3), and 23–34% based on the 2SLS estimates in Column (4). The remaining share is attributable to variation in the strength of knowledge diffusion across distance bins.

Table 2: Impact of Education and Knowledge Diffusion on Innovation Productivity

	Num. of self-inv.	Num. of self-inv. per 100 R&D workers				
	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Distance bin 1	0.894***	0.562***	0.169	0.104	0.210*	0.142
	(0.199)	(0.126)	(0.122)	(0.127)	(0.120)	(0.125)
Distance bin 2	0.609***	0.586***	0.274*	0.223	0.305**	0.252
	(0.107)	(0.155)	(0.146)	(0.153)	(0.149)	(0.156)
Distance bin 3	0.711***	0.254***	0.0992*	0.0719	0.114**	0.0850
	(0.228)	(0.0551)	(0.0540)	(0.0567)	(0.0541)	(0.0570)
Distance bin 4	0.117	0.109**	0.0205	0.0134	0.0291	0.0203
	(0.0734)	(0.0535)	(0.0527)	(0.0528)	(0.0528)	(0.0531)
Share of univ. graduates			0.988***	2.517***	0.822***	2.516***
			(0.132)	(0.873)	(0.0969)	(0.884)
× new entrant					0.321	−0.160
					(0.208)	(0.482)
Knowledge diffusion			0.974***	0.979***	0.835***	0.829***
			(0.0283)	(0.0315)	(0.0313)	(0.0351)
× new entrant					0.404***	0.425***
					(0.0682)	(0.0759)
New entrant					−0.223*	0.0358
					(0.120)	(0.338)
Observations	165,722	165,722	165,722	165,717	165,722	165,717
R-squared	0.011	0.004	0.009	0.005	0.010	0.005
First-stage F				722.35		358.99

Notes: OLS and 2SLS regressions of firm-level innovation outcomes on distance-to-university bins, the share of university graduates among R&D personnel, and the knowledge diffusion measure defined in the text. Column (1) uses total self-invention patents as the dependent variable; Columns (2)–(6) use self-invention patents per 100 R&D workers. Columns (5)–(6) add interactions with a new-entrant dummy (firm age < 9 years). Instruments in Columns (4) and (6) are constructed as in equation (3), excluding universities within 10 km. All regressions include city, year, and industry fixed effects, as well as sales decile dummies. First-stage F refers to the Kleibergen-Paap statistic. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

New firms' responses. We next examine if the age of firms affects the documented relationship between firm innovation intensity and proximity to universities, considering that new firms are typically more innovative and may also face difficulties in obtaining innovation resources. A firm is considered new if it is less than 9 years old, following the categorization used by Acemoglu, Akcigit, Alp, Bloom & Kerr (2018). We add the interaction between the new firm dummy and the two channels into regression (2) and accordingly construct instruments for these interaction terms using the interaction between instruments in equation (3) and the new firm dummy. The OLS and IV results in Columns (5) and (6) of Table 2 show that new firms can benefit more from universities, especially as they have higher efficiency in absorbing knowledge diffusion.

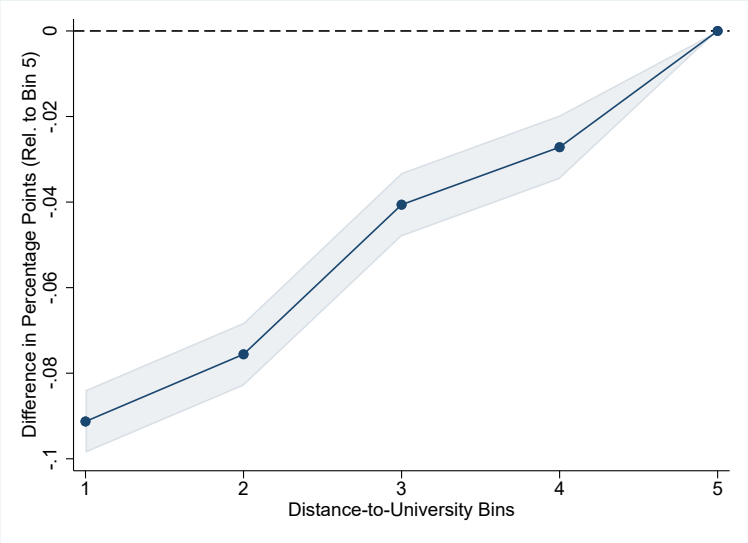
Are new firms able to reap the benefits from proximity to universities by locating near them? To investigate this, we estimate a linear probability model to examine the likelihood of a firm choosing to locate close to a university:

$$new_{i,t} = \beta_0 + \sum_{r=1}^{4} \beta_r \times bin_{r(i,t)} + \phi \mathbf{X}_{it} + \delta_{c(i,t)} + \eta_t + \epsilon_{it}. \tag{4}$$

β_r measures the relative probability that new firms choose to locate in bin r compared with the farthest bin ($r = 5$, the baseline group). As in previous specifications, we include firm-level controls, city fixed effects, and time fixed effects. Firms are divided equally into five distance bins within each city and year; if new firms chose locations with the same distribution as incumbents, we would expect $\beta_r = 0$ for all r.

Figure 3 presents the results, showing that new firms are relatively less likely than incumbents to locate near universities. This pattern suggests the presence of constraints for new firms to establish themselves close to universities. One natural and key constraint is likely land availability: locating near universities requires offices or factories in those areas, and because incumbents occupy these locations, new firms may face difficulties finding suitable space.

Figure 3: Probability of Entry of New Firms (relative to last bin)



Notes: Estimated coefficients $\hat{\beta}_r$ from equation (4), a linear probability model with the new-firm indicator (age < 9 years) as the dependent variable. Within each city-year, firms are divided equally into five distance bins; bin 5 (farthest) is the omitted category. All regressions include firm-level controls, city fixed effects, and year fixed effects. Dotted lines represent 95% confidence intervals based on robust standard errors.

Robustness checks. We conduct several tests to verify the robustness of the empirical results reported in Table 2. First, because the number of patents is count data, we re-estimate the model using a Poisson pseudo-maximum likelihood (PPML) estimator instead of ordinary least squares. The results, presented in Appendix Table A.2, confirm the robustness of our findings. Here, the coefficients should be interpreted as the proportional change in patent counts associated with a one-unit change in each explanatory variable.

We also test several alternative instruments. First, in constructing industry-specific reliance on high-skilled talent (equation 3), we originally used $SI^{-i}_{s(i),t}$, the nationwide industry-level share of university graduates in R&D, excluding the firm itself. To address potential spatially correlated R&D shocks (e.g., new innovation hubs), which may affect local demand for high-skilled talent, we re-compute this measure excluding same-city firms, $SI^{-c}_{s(i),t}$. Second, instead of excluding universities within 10 km of each firm, we also consider 5 km and 20 km thresholds in constructing the set of excluded universities $\mathcal{C}_i$. As shown in Appendix Table A.3, our IV results in Table 2 remain robust to these alternative constructions.

2.3 Location Choices of College Graduates

Finally, we use Revelio Labs data to examine whether university graduates tend to remain near their alma maters, which would reinforce the talent-pool channel documented above. We geocode company and university addresses and retain only firms and universities with available location information for individuals who received their college education in China. Each observation in the dataset represents a job, corresponding to an individual holding a specific position at a firm. We restrict the sample to jobs with start dates no earlier than the individual's graduation date. We obtain information for roughly 2.9 million jobs. Although the data may be subject to selection issues, it represents the largest available employer–employee dataset for China.

First, Table 3 shows that graduates are more likely to remain in the city where they attended college: 31.4% of graduates stay in their college city, compared with only 8.8% under a random nationwide distribution. Second, even within the same city, local university graduates tend to work closer to universities than non-local graduates. Figure 4 displays the distance distribution, and Table 3 indicates that, conditional on city (city-sector) and year, local graduates are on average 11.5% (5.4%) closer to universities than their non-local counterparts.

Table 3: Proximity of Local College Graduates to Universities

Panel A: Probability of working in the city of college attendance

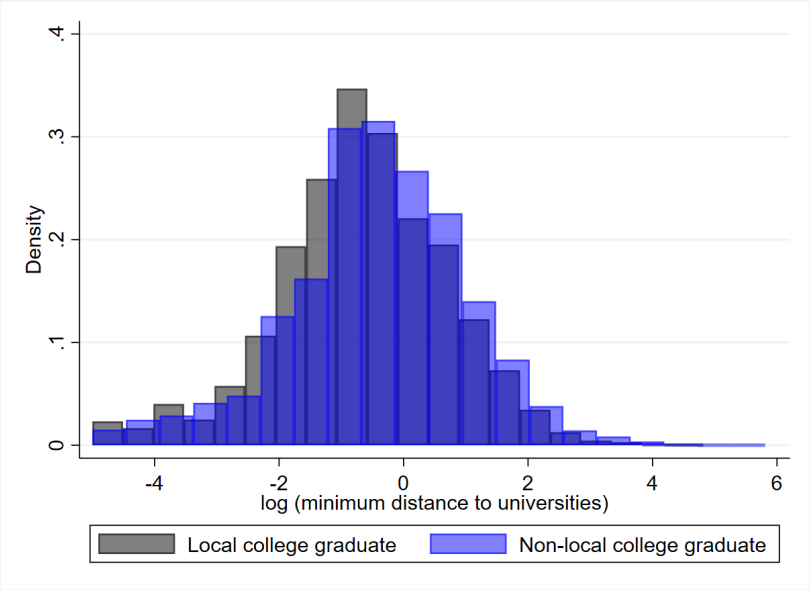
	Work in city of college attendance	Leave city of college attendance
Actual data	31.4%	68.6%
Random assignment (hypothetical)	8.8%	91.2%

Panel B: Proximity to universities within the city

Minimum distance to universities (local graduates vs. non-local)		
Raw data	Conditional on same city and year	Conditional on same city-sector and year
−24.0%	−11.5%	−5.4%

Notes: Panel A compares the probability that a graduate works in the city where they attended college against a hypothetical random assignment across cities. Panel B reports the percentage difference in minimum distance to universities between local graduates (those working in their college city) and non-local graduates. "Raw data" is unconditional; remaining columns condition on city and year, or city-sector and year fixed effects. Data from Revelio Labs; sample restricted to individuals with Chinese college education and jobs starting after graduation (≈2.9 million observations).

Figure 4: Distribution of Distance to the Nearest Universities



Notes: Histogram of the log minimum distance (in km) from each worker's employer to the nearest university, separately for local college graduates (those working in the city where they attended college) and non-local college graduates. Data from Revelio Labs; sample restricted to individuals with Chinese college education and jobs starting after graduation.

3 Model: Innovation, Land, and Location Choices

Innovation models have traditionally paid little attention to geography. However, our evidence shows that proximity to universities plays an important role: firms located nearby produce more innovations, while entrants—who typically choose locations farther away—do not benefit from this advantage. This generates a tradeoff in which firms compete in product space independently of their location choices. As a result, when one firm innovates and displaces another, it does not inherit the incumbents favorable location, potentially leading to spatial misallocation in innovation. In addition, because skilled workers tend to remain near the universities from which they graduate, the concentration of talent further reinforces the benefits of locating close to universities.

3.1 Simple Model

Innovation models have traditionally paid little attention to geography. However, our evidence reveals that proximity to universities matters: firms located nearby innovate more

efficiently, while new entrants—who are more productive innovators—tend to locate farther away. This pattern arises because incumbents occupy scarce land near universities, pushing entrants to less productive locations. Crucially, when an entrant innovates and displaces an incumbent, it takes over the incumbent's *product line* but *not* the incumbent's physical location. This decoupling between product-space competition and physical-space allocation is a key source of spatial misallocation that standard growth models miss entirely.

To formalize these ideas, we extend the canonical Schumpeterian framework to include two locations that differ in innovation productivity and land availability. The model highlights three forces: (i) knowledge spillovers from universities raise innovation productivity in nearby locations; (ii) limited land near universities constrains how many firms can locate there; and (iii) creative destruction determines land availability, but freed land does not follow the innovator. We show that the resulting equilibrium features spatial misallocation—too few innovators near universities relative to the social optimum—and that land constraints amplify the gap between market and socially optimal innovation.

Environment. Consider an economy with two locations, $d \in \{1, 2\}$, interpreted as *near* and *far* from a university. Each location has an exogenous innovation productivity A_d, with $A_1 \geq A_2$, reflecting knowledge spillovers that decay with distance. The total land in the economy is $L = L_1 + L_2$, where L_d denotes the land supply at location d. There are two types of workers—production workers with supply N^u and R&D workers with supply N^h—each trading in a separate competitive labor market. All varieties are freely traded across locations.

Final-good producers. A unique final consumption good is produced competitively by assembling all varieties:

$$Y_t = \frac{1}{1-\beta} \int_0^1 q_{j,t}^{\beta} k_{j,t}^{1-\beta} dj, \tag{5}$$

where $q_{j,t}$ is the quality of variety j and $k_{j,t}$ is the quantity demanded. The inverse demand function is $p_{j,t} = q_{j,t}^{\beta} k_{j,t}^{-\beta}$.

Producers of intermediate goods. Intermediate good producers require land and production labor to produce each variety under monopolistic competition. A firm producing variety j at location d with land l and production labor u (in efficiency units) produces

$$k_{j,t} = \bar{q}_t u l^{\theta},$$

where $\bar{q}_t \coloneqq \int_0^1 q_{j,t} dj$ is aggregate quality and $\theta \geq 0$ governs the importance of land. Since varieties are freely traded, the profit-maximizing labor choice and the resulting profits are location-independent conditional on quality and land. As before, the optimal profit for a firm with quality q_j and land l is

$$\pi_{j,t} = \pi_t \, q_j \, l^{\alpha_l}, \tag{6}$$

where $\pi_t = \frac{\beta}{1-\beta}(1-\beta)^{\frac{1}{\beta}} \left(\frac{w_t^u}{\bar{q}_t}\right)^{\frac{\beta-1}{\beta}}$ and $\alpha_l = \frac{\theta(1-\beta)}{\beta}$.

Entrants: innovation and location choice. A unit mass of identical potential entrants choose both a location $d \in \{1, 2\}$ and their innovation investment. Let $\xi \in [0, 1]$ denote the share of entrants choosing $d = 1$. At location d, an entrant hires R&D labor h (at wage w_t^h) and acquires land l to maximize expected profits. Innovation is *undirected*: a successful innovator at location d takes over a randomly drawn product line from the unit interval, improving its quality by $\lambda \bar{q}_t$.

Key spatial friction. When an entrant at location d innovates on product j, the previous incumbent of j exits and frees land at the *incumbent's* location—which may differ from d. The new firm operates product j from location d, not from the displaced firm's location. Innovation reallocates product lines across locations but does not transfer physical space.

After taking over a product, the entrant collects monopoly profits until creatively destroyed at rate x. The value function of an incumbent firm satisfies the HJB equation

$$r_t V_t(q_j, l) - \dot{V}_t(q_j, l) = \pi_t \, l^{\alpha_l} \, q_j - x_t V_t(q_j, l). \tag{7}$$

Since profits are location-independent conditional on (q_j, l), the value function takes the

form $V_t(q_j, l) = v_t\, q_j\, l^{\alpha_l}$.

At location d, an entrant takes the land price $\kappa_{d,t}$ and the R&D wage w_t^h as given and solves

$$\max_{l,h}\ \phi_d(h)\, v_t\, \bar{q}_t(1+\lambda)\, l^{\alpha_l} - w_t^h\, h - \kappa_{d,t}\, l,$$

where $\phi_d(h) = A_d\, h^{\alpha_h}$ is the innovation production function at location d, with A_d capturing university proximity effects and $\alpha_h \in (0,1)$. The first-order conditions are

$$[h]:\ \alpha_h\, A_d\, h^{\alpha_h - 1}\, v_t\, \bar{q}_t(1+\lambda)\, l^{\alpha_l} = w_t^h, \tag{8}$$

$$[l]:\ \phi_d(h)\, \alpha_l\, v_t\, \bar{q}_t(1+\lambda)\, l^{\alpha_l - 1} = \kappa_{d,t}. \tag{9}$$

Location choice. If both locations are active ($\xi \in (0,1)$), entrants must be indifferent between them. The maximized expected payoff at location d is $(1-\alpha_h-\alpha_l)\,\phi_d\, v_t\, \bar{q}_t(1+\lambda)\, l_d^{\alpha_l}$.[8] Indifference therefore requires

$$\phi_1\, l_1^{\alpha_l} = \phi_2\, l_2^{\alpha_l}. \tag{10}$$

Aggregate creative destruction equals total innovation across both locations:

$$x = \xi\, \phi_1 + (1-\xi)\, \phi_2. \tag{11}$$

Land market. At each location d, land is freed when local incumbents are displaced through creative destruction. Since innovation is undirected, every product line faces the same displacement rate x, so the flow of available land at d is $x\, L_d$. The land market clearing condition at each location equates entrant demand to available supply:

$$\xi\, l_1 = x\, L_1, \qquad (1-\xi)\, l_2 = x\, L_2. \tag{12}$$

Importantly, land freed by an exiting firm at location d is available *only* to entrants who choose d, it does not follow the innovator to a different location. This decoupling between product and physical spaces leads to spatial misallocation.

[8] This follows from substituting the FOCs back into the objective. The SOC requires $\alpha_h + \alpha_l < 1$.

Workers. Workers maximize the expected value of log utility of consuming the final good and discount the future at rate ρ. They supply labor inelastically and the Euler equation implies $r_t = \rho + \dot{c}_t/c_t$.

Labor markets. The two labor markets clear separately. Let μ_d denote the share of product lines operating from location d on the balanced growth path. Production labor demand aggregates across all intermediate producers:

$$\left(\frac{(1-\beta)\bar{q}_t}{w_t^u}\right)^{1/\beta} \left[\mu_1 \, l_1^{\alpha_l} + \mu_2 \, l_2^{\alpha_l}\right] = N^u. \tag{13}$$

R&D labor demand aggregates across all entrants:

$$\xi \, h_1 + (1-\xi) \, h_2 = N^h. \tag{14}$$

Equilibrium. A balanced growth path (BGP) equilibrium consists of allocations $\{k_{j,t}^*, u_{j,t}^*, h_d^* \, , l_d^* \, , c_t^*\}$, prices $\{p_{j,t}, w_t^u, w_t^h, r_t, \kappa_{d,t}\}$, a location allocation ξ^*, and the value function $V_t(q,l)$ such that all agents optimize, the value function satisfies (7), and all markets clear.

Balanced growth. On a BGP, aggregate quality grows at a constant rate $g = x\lambda$. Land per firm l_d, R&D labor h_d, production labor allocations, and the interest rate are constant over time. Wages, land prices, and aggregate output grow at rate g. The value function is stationary with $v = \pi^*/(r^* + x^*)$.

A key property of the equilibrium is that the indifference condition (10) implies $A_1 l_1^{\alpha_l} = A_2 l_2^{\alpha_l}$, which in turn equalizes R&D labor across locations: $h_1^* = h_2^*$.[9] Together with the R&D labor market clearing (14), this gives $h_1^* = h_2^* = N^h$, and the equilibrium innovation rates are $\phi_d^* = A_d (N^h)^{\alpha_h}$. Substituting into the land market clearing conditions (12) yields the following characterization. All proofs are in Appendix B.1.

Proposition 1 (Market Equilibrium)**.** *Suppose $\alpha_h + \alpha_l < 1$ and $A_1 \geq A_2 > 0$. There exists a unique interior BGP equilibrium (ξ^*, x^*) with both locations active. The equilibrium location*

[9] Since the marginal return to R&D labor depends on $A_d l_d^{\alpha_l}$ (from the FOC (8)), equalization of this product implies identical optimal R&D hiring at both locations.

allocation satisfies

$$\frac{A_1}{A_2} = \left(\frac{\xi^* L_2}{(1-\xi^*) L_1} \right)^{\alpha_l}, \tag{15}$$

and the creative destruction rate is $x^* = (N^h)^{\alpha_h}[\xi^* A_1 + (1-\xi^*)A_2]$.

The equilibrium has the following properties:

(i) ***Sorting:*** $\xi^* > L_1/L$ *if and only if* $A_1 > A_2$. *More entrants locate near the university than the land share alone would predict.*

(ii) ***Land–productivity tradeoff:*** *Firms near the university hold less land* ($l_1^* < l_2^*$), *hire the same R&D labor* ($h_1^* = h_2^* = N^h$), *and innovate at a higher rate* ($\phi_1^* > \phi_2^*$) *purely due to higher* A_1.

Condition (15) captures the central tradeoff facing entrants: locating near the university raises innovation productivity ($A_1 > A_2$) but increases competition for scarce land, reducing l_1 and hence production scale. In equilibrium, higher land costs near the university exactly offset the innovation advantage. Notably, the location allocation (15) does not depend on α_h: the elasticity of innovation with respect to R&D labor cancels in the indifference condition, so the spatial equilibrium is governed entirely by the interplay between innovation productivity and land costs.

Social Planner's Problem

Consider a social planner who maximizes worker utility subject to the resource and technological constraints, with the additional choice over the spatial allocation of innovation:

$$\begin{aligned}
\max_{\xi, h_d, l_d, \{u_j\}_j} \quad & \int_0^\infty e^{-\rho t} \log C_t \, dt \\
\text{s.t. } & Y_t = C_t, \quad Y_t = \frac{1}{1-\beta} \int q_j^\beta \bar{q}_t^{1-\beta} u_{jt}^{1-\beta} l_t^{(1-\beta)\theta} dj \\
& \frac{\dot{\bar{q}}_t}{\bar{q}_t} = \left[\xi \, \phi_1(h_1) + (1-\xi) \, \phi_2(h_2)\right] \lambda, \\
\text{Labor:} \quad & \xi \, h_1 + (1-\xi) \, h_2 = N^h, \quad \int u_j dj = N^u \\
\text{Land:} \quad & \xi \, l_1 = x \, L_1, \quad (1-\xi) \, l_2 = x \, L_2.
\end{aligned}$$

As in the market equilibrium, we focus on the BGP where aggregate quality grows at a constant rate $g = x\lambda$ and all allocations (ξ, h_d, l_d) are constant over time. On a BGP, the planner's objective reduces to maximizing a weighted sum of the growth rate and the level of initial output (see Appendix B.1 for details). The planner therefore faces a tradeoff between concentrating innovation near universities to raise growth versus spreading firms across locations to increase output through better land utilization.

Compared with the market equilibrium, the planner has two additional margins of adjustment. First, the planner can allocate R&D labor *differentially* across locations: while the market equalizes R&D labor ($h_1^* = h_2^*$), the planner tilts R&D effort toward the more productive location. Second, the planner can adjust the spatial allocation ξ to internalize the growth externality from creative destruction.

Proposition 2 (Social Planner)**.** *Under the conditions of Proposition 1, there exists a unique BGP solution (ξ^s, h_1^s, h_2^s) to the planner's problem, with associated creative destruction rate x^s and growth rate $g^s = x^s\lambda$. The planner's R&D labor allocation satisfies*

$$\frac{h_1^s}{h_2^s} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\alpha_h}} > 1 \quad \textit{when} \quad A_1 > A_2. \tag{16}$$

The planner allocates more R&D labor to the near-university location because the social

return to innovation—its contribution to aggregate growth—is proportional to A_d. In the market equilibrium, by contrast, the private return to R&D is mediated by land: each entrant's payoff depends on $A_d l_d^{\alpha_l}$, which is equalized across locations on the BGP. The planner sees through this mediation, recognizing that the growth externality scales with A_d alone.

By comparing the market and planner's BGP allocations, we characterize the spatial inefficiency.

Proposition 3 (Spatial Inefficiency)**.** *Suppose $\alpha_h + \alpha_l < 1$ and $A_1 > A_2$. On the BGP:*

(i) ***R&D labor misallocation:*** *The market equalizes R&D labor ($h_1^* = h_2^*$) while the planner concentrates it near the university ($h_1^s > h_2^s$), yielding strictly higher creative destruction: $x^s > x^*$.*

(ii) ***Location misallocation:*** *The planner concentrates a weakly larger share of entrants near the university: $\xi^s \geq \xi^*$, with the gap increasing in A_1/A_2.*

(iii) ***Innovation gap amplification:*** *The planner's innovation rate ratio exceeds the market's: $\phi_1^s/\phi_2^s = (A_1/A_2)^{1/(1-\alpha_h)} > A_1/A_2 = \phi_1^*/\phi_2^*$.*

Part (i) is the novel spatial result. The market cannot channel R&D effort to where it is most productive because private returns are mediated by the land market: the indifference condition equalizes the product $A_d l_d^{\alpha_l}$ across locations, forcing R&D labor to be equalized even though the social return to R&D is higher near the university. The planner, who maximizes aggregate creative destruction for a given labor budget, allocates more R&D workers to $d = 1$, achieving strictly higher growth. Part (ii) reinforces this: the planner also sends more entrants to the productive location, accepting greater land crowding for faster growth. Part (iii) shows that these two margins compound, so that the planner's innovation gap between locations exceeds the market's.

This spatial dimension of inefficiency goes beyond the standard under-investment result in Schumpeterian models. Not only is overall innovation below the social optimum, but both its *geographic composition* (where firms locate) and its *intensity composition* (how much R&D effort they exert) are distorted.

Discussion. The severity of spatial misallocation depends on the strength of university spillovers (A_1/A_2) and the importance of land in production (θ). Stronger spillovers widen the gap between private and social returns to locating near universities, as the market undervalues the growth externality from high-productivity innovation. Higher θ raises α_l, making firms more sensitive to land scarcity and pushing more entrants to the distant location, amplifying the spatial distortion. Our counterfactual exercises in Section 4 quantify these effects.

Having established the theoretical mechanisms in this simplified setting, we now extend the model to a fully quantitative framework with multiple regions and distance bins to measure the magnitude of these spatial inefficiencies.

3.2 Quantitative Model

In this section, we outline how the simple model is extended into a quantitative framework by incorporating multiple regions and firms' distances to universities. This extension is intended to capture our empirical findings on how the joint location decisions of firms and workers shape aggregate innovation outcomes and growth. To avoid redundancy, we discuss only the components that depart from the simple model.

Consider an economy with N regions, each partitioned into D discrete bins based on their distance to universities, $d = 1, 2, \ldots, D$. Each region-bin hosts a continuum of firms, each producing a distinct variety. The measure of firms in each location is determined endogenously through entry and creative destruction, as outlined below. We refer to each region-bin (n, d), for $n = 1, \ldots, N$ and $d = 1, \ldots, D$, as a location r, where production takes place. Varieties are freely traded across all locations. Unlike in the simple model, labor markets are now segmented both by skill type and by location.

Innovation by incumbent firms. The simple model assumed that only entrants innovate. Here, incumbent producers also innovate on their products by hiring R&D labor to improve the quality of their current intermediate goods. To reflect the lifecycle pattern in innovation intensities, we posit that incumbent firms in location r start with innovation efficiency A_r. However, over time, they can transition to a lower innovation efficiency, ζA_r, with

$\zeta \leq 1$, with probability ψ. We refer to firms with innovation efficiency A_r as *high-innovation-capacity* firms and firms with innovation efficiency ζA_r as *low-innovation-capacity* firms. As in the simple model, if innovation is successful, the quality of the product increases by step size $\lambda \bar{q}_t$, depending on the economy's aggregate quality $\bar{q}_t$.

Incumbent firms are replaced by others at the creative destruction rate x. This rate is determined endogenously in equilibrium and influences not only the innovation rate in the typical manner highlighted in the innovation literature but also the availability of land for new entrant firms.

Production. The simple model used a single type of production labor u. Here, the production function combines skilled workers s and unskilled workers u via a CES aggregator:

$$k = \bar{q}_t \left(\Psi u^{\frac{\rho-1}{\rho}} + (1-\Psi) s^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} l^{\theta},$$

where $\rho > 1$ is the elasticity of substitution between skilled and unskilled workers, Ψ governs the efficiency of unskilled labor in production, and θ is the elasticity of production with respect to land, as before. Denoting the wages for skilled and unskilled workers in location r by $w^s_{r,t}$ and $w^u_{r,t}$, the composite wage cost is

$$w_{r,t} = \left(\Psi^{\rho} \left(w^u_{r,t}\right)^{1-\rho} + (1-\Psi)^{\rho} \left(w^s_{r,t}\right)^{1-\rho} \right)^{1/(1-\rho)}.$$

The firm producing variety j in location r maximizes profits $p_{j,t} k_{j,t} - \frac{w_{r,t}}{\bar{q}_t l^{\theta}} k_{j,t}$, yielding optimal output and profits:

$$k_{j,t} = \left[\frac{(1-\beta)\bar{q}_t}{w_{r,t}} \right]^{\frac{1}{\beta}} q_{j,t} l^{\frac{\theta}{\beta}}, \qquad \pi_{j,t} = \pi_{r,t} \, l^{\alpha_l} \, q_{j,t},$$

where $\pi_{r,t} = \beta \left[\frac{(1-\beta)\bar{q}_t}{w_{r,t}} \right]^{\frac{1-\beta}{\beta}}$ and $\alpha_l = \frac{\theta(1-\beta)}{\beta}$ as before.

R&D labor. The R&D labor input is also a CES composite of skilled and unskilled labor:

$$h = \left(\Psi^e u^{\frac{\rho-1}{\rho}} + (1-\Psi^e) s^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}},$$

where Ψ^e governs the productivity of unskilled labor in R&D. The unit cost of R&D labor in location r is

$$w^e_{r,t} = \left[(\Psi^e)^\rho (w^u_{r,t})^{1-\rho} + (1-\Psi^e)^\rho (w^s_{r,t})^{1-\rho}\right]^{\frac{1}{1-\rho}}.$$

This replaces the single R&D wage w^h_t from the simple model.

Incumbent firms' problem. An incumbent firm chooses R&D labor h to maximize the discounted value of expected lifetime profits. Incumbents differ in their location, innovation capacity, land holdings, and product quality. The value function of a high-innovation-capacity firm with quality q_j and land l in location r is:

$$\begin{aligned} r_t V^H_{r,t}(q_j,l) - \dot{V}^H_{r,t}(q_j,l) = \max_h \; & \pi_{r,t}\, q_j\, l^{\alpha_l} + \phi_r(h)\left[V^H_{r,t}(q_j+\lambda\bar{q}_t,l) - V^H_{r,t}(q_j,l)\right] \\ & - (w^e_{r,t}h + \xi\bar{q}_t h^\gamma) + \psi\left[V^L_{r,t}(q_j,l) - V^H_{r,t}(q_j,l)\right] - x V^H_{r,t}(q_j,l), \end{aligned}$$

where $\phi_r(h) = A_r h^{\alpha_h}$ is the innovation production function, with A_r capturing location-specific R&D productivity as in the simple model. We also assume a convex cost for hiring R&D labor, $\xi\bar{q}_t h^\gamma$, with $\xi \geq 0$ and $\gamma > 1$. This reflects the idea that the marginal cost of hiring R&D workers increases with the size of R&D personnel, which can be rationalized by search and matching frictions in the R&D labor market with convex vacancy costs (Acemoglu & Hawkins 2014), or by the difficulty of screening workers' abilities (Helpman, Itskhoki, Muendler & Redding 2016). Quantitatively, without such convex costs, differences in innovation productivity due to proximity to universities would be driven solely by land prices, making it difficult to match the research productivity advantage of firms closer to universities documented in Section 2.

The first-order condition equates the marginal value of R&D labor to its marginal cost:

$$[h]: \;\; \alpha_h A_r h^{\alpha_h - 1}\left[V^H_{r,t}(q_j+\lambda\bar{q}_t,l) - V^H_{r,t}(q_j,l)\right] = w^e_{r,t} + \xi\gamma\bar{q}_t h^{\gamma-1}. \tag{17}$$

We denote the optimal R&D labor as $h^H_{r,t}(q_j,l)$ and the corresponding innovation rate as $\phi^H_{r,t}(q_j,l) = A_r\left(h^H_{r,t}(q_j,l)\right)^{\alpha_h}$.

Similarly, the value function of a low-innovation-capacity firm with quality q_j and land l in

location r is:

$$r_t V^L_{r,t}(q_j, l) - \dot{V}^L_{r,t}(q_j, l) = \max_h \ \pi_{r,t}\, q_j\, l^{\alpha_l} + \zeta\phi_r(h)\left[V^L_{r,t}(q_j + \lambda\bar{q}_t, l) - V^L_{r,t}(q_j, l)\right] - \left(w^e_{r,t} h + \xi\bar{q}_t h^\gamma\right) - x V^L_{r,t}(q_j, l).$$

The main difference relative to high-capacity firms is the reduced innovation productivity $\zeta \leq 1$. The FOC is:

$$[h]: \ \alpha_h \zeta A_r h^{\alpha_h - 1}\left[V^L_{r,t}(q_j + \lambda\bar{q}_t, l) - V^L_{r,t}(q_j, l)\right] = w^e_{r,t} + \xi\gamma\bar{q}_t h^{\gamma-1}. \tag{18}$$

Let $h^L_{r,t}(q_j, l)$ denote the optimal R&D labor and $\phi^L_{r,t} = \zeta A_r\left(h^L_{r,t}(q_j, l)\right)^{\alpha_h}$ the corresponding innovation rate.

Innovation by entrants. In each location r, potential entrants face a fixed entry cost of $\eta^e_r\bar{q}_t$ in units of the final good, capturing registration and setup expenses. Upon entry, firms choose their innovation rate. If the innovation is successful, the firm also chooses the amount of land l, which remains fixed throughout the firm's lifetime until exit. A successful innovation randomly replaces an existing variety and delivers a quality improvement of size $\lambda\bar{q}_t$. Firms enter as high-innovation-capacity producers, although their innovation capabilities may decline later.

We assume that the land market is competitive, so entrants take the price of land $\kappa_{r,t}$ in location r as given and solve

$$V^e_{r,t} = \max_{\{h,l\}} \phi_r(h)\left[\mathbb{E}V^H_{r,t}(q_j + \lambda\bar{q}_t, l) - \kappa_{r,t} l\right] - \left(w^e_{r,t} h + \xi\bar{q}_t h^\gamma\right),$$

where the expectation captures uncertainty regarding which variety will be replaced. For simplicity, we assume that all successful innovators select the land area l before the realization of the specific variety they replace, ensuring that all successful innovations choose the same amount of land.

The first-order conditions are

$$[h]: \ \alpha_h A_r h^{\alpha_h - 1}\left[\mathbb{E}V^H_{r,t}(q_j + \lambda\bar{q}_t, l) - \kappa_{r,t} l\right] = w^e_{r,t} + \xi\gamma\bar{q}_t h^{\gamma-1},$$
$$[l]: \ \frac{\partial \mathbb{E}V_{r,t}(q_j + \lambda\bar{q}_t, l)}{\partial l} = \kappa_{r,t}.$$

In equilibrium, free entry requires:

$$V^e_{r,t} \leq \eta^e_r \bar{q}_t,$$

with equality if the number of entrants $M^e_{r,t}$ is strictly positive. We denote the optimal choices as $l^e_{r,t}$ and $h^e_{r,t}$, with the corresponding innovation rate $\phi^e_{r,t}$.

Workers. We denote the measure of unskilled workers in location r at time t by $N^u_{r,t}$. For skilled workers, we use $N^{s(c)}_{r,t}$, where c indexes the region in which the worker attended college. At the end of each period, a random fraction δ of workers retire. The remaining workers and young entrants choose whether to migrate.

We build on the frameworks of Artuc, Chaudhuri & McLaren (2010) and Caliendo, Dvorkin & Parro (2019) to model workers' dynamic migration decisions, with particular attention to the continuous-time formulation presented in Bilal & Rossi-Hansberg (2023). Specifically, we assume that at each point in time, the opportunity to move arrives with a Poisson rate μ, in which case workers draw idiosyncratic location preference shocks $\{\epsilon_r\}$ according to a Type-I Extreme Value distribution $F_\epsilon(\epsilon) = \exp\left(-\exp\left(-\frac{\epsilon}{\nu} - \gamma\right)\right)$, i.i.d. over time and across locations, with γ denoting Euler's constant. The parameter $\nu > 0$ governs the dispersion of location preferences. The migration cost from location r to r' is $d^\tau_{r,r'} \geq 0$.

Define the value function of a worker of skill type $\tau \in \{u, s(c)\}$ in location r at time t as $J^\tau_{r,t}$. The worker's HJB equation is:

$$\rho J^\tau_{r,t} - \dot{J}^\tau_{r,t} = B^\tau_r + \log(C^\tau_{r,t}) + \mu\mathcal{M}^\tau_r\left(\mathbf{J}^\tau_t\right) - \delta J^\tau_{r,t},$$

where B^τ_r denotes amenities and $\mathcal{M}^\tau_r\left(\mathbf{J}^\tau_t\right)$ represents the option value of migration:

$$\mathcal{M}^\tau_r(\mathbf{J}^\tau_t) = \nu \log\left(\sum_{r'} \exp\left(J^\tau_{r',t} - d^\tau_{r,r'}\right)^{1/\nu}\right) - J^\tau_{r,t}.$$

Among workers permitted to relocate from r, the share choosing r' is:

$$\Pi^{\tau}_{r,r',t} = \frac{\exp\left(J^{\tau}_{r',t} - d^{\tau}_{r,r'}\right)^{1/\nu}}{\sum_{r'} \exp\left(J^{\tau}_{r',t} - d^{\tau}_{r,r'}\right)^{1/\nu}}.$$

The change in employment in location r is:

$$\frac{dN^{\tau}_{r,t}}{dt} = \sum_{r'} \Pi^{\tau}_{r',r,t}\left(\mu N^{\tau}_{r',t} + N^{\tau,e}_{r',t}\right) - (\mu + \delta) N^{\tau}_{r,t},$$

where $N^{\tau,e}_{r,t}$ denotes the measure of new young workers entering location r. For unskilled workers, we assume $N^{u,e}_{r,t} = \delta L^{u}_{r,t}$, so retired workers are replaced by offspring in the same location. For skilled workers, we assume $N^{s(r),e}_{r',t} = 0$ for all $r \neq r'$, with $N^{s(r),e}_{r,t}$ determined by the data on college graduates in region r. Under this assumption, college graduates enter the labor market in the region where they attended college, but may relocate immediately upon entering the workforce.

Labor market clearing. Denote the measures of active high- and low-innovation-capacity firms as $M^{H}_{r,t}$ and $M^{L}_{r,t}$, respectively. The labor market clears in each location, separately for unskilled and skilled workers. For unskilled workers:

$$\underbrace{\int_{j \in r} \frac{k_{j,t}}{\bar{q}_t l^{\theta}} dj \cdot \frac{\Psi^{\rho}\left(w^{u}_{r,t}\right)^{-\rho}}{w^{-\rho}_{r,t}}}_{\text{production demand}} + \underbrace{\left[M^{H}_{r,t} h^{H}_{r,t} + M^{L}_{r,t} h^{L}_{r,t} + M^{e}_{r,t} h^{e}_{r,t}\right] \cdot \frac{\left(\Psi^{e}\right)^{\rho}\left(w^{u}_{r,t}\right)^{-\rho}}{\left(w^{e}_{r,t}\right)^{-\rho}}}_{\text{R\&D demand}} = N^{u}_{r,t}.$$

An analogous condition holds for skilled workers, with Ψ and Ψ^{e} replaced by $(1 - \Psi)$ and $(1 - \Psi^{e})$, and supply given by $\sum_{c} N^{s(c)}_{r,t}$.

Land market clearing. For each location r, there is a total supply of land L_r owned by exogenous landowners who spend all their income on final goods. As in the simple model, the available land depends on the rate of creative destruction x:

$$M^{e}_{r,t} \phi^{e}_{r,t} l^{e}_{r,t} = x L_r, \quad \forall\, r.$$

Balanced Growth Path

We focus on characterizing the equilibrium on a balanced growth path (BGP), where aggregate productivity $\bar{q}_t$ grows at a constant rate $g := \dot{\bar{q}}_t/\bar{q}_t$. From the final-good production technology (5), this implies the overall economy grows at the same rate g.

From the intermediate-good producers' problem, wages $w^u_{r,t}$ and $w^s_{r,t}$ must grow at rate g, implying that the profit term $\pi_{r,t} = \pi_r$ is constant.

On a BGP, R&D labor $h^\tau_r(q_j, l)$, with $\tau \in \{H, L\}$, is constant and does not depend on q_j, as shown below. We suppress its dependence on land l, since all firms within location r choose the same land area. The innovation arrival rate is therefore constant: $\phi^\tau_r = A_r\big(h^\tau_r\big)^{\alpha_h}$.

We conjecture that the value functions on the BGP take the form $V^\tau_{r,t}(q_j, l) = (v_r q_j + \bar{v}^\tau_r \bar{q}_t) l^{\alpha_l}$, where v_r and $\bar{v}^\tau_r$ are scalars. Substituting into the HJB equations yields:

$$v_r = \frac{\pi_r}{\rho + x}, \qquad \bar{v}^L_r = \frac{(1-\alpha_h)\phi^L_r v_r \lambda + \xi(\gamma-1)(h^L_r)^\gamma l^{-\alpha_l}}{\rho - g + x},$$
$$\bar{v}^H_r = \frac{(1-\alpha_h)\phi^H_r v_r \lambda + \xi(\gamma-1)(h^H_r)^\gamma l^{-\alpha_l} + \psi \bar{v}^L_r}{\rho - g + x + \psi}.$$

Substituting these into the FOCs (17) and (18) confirms that the optimal R&D labor h^H_r and h^L_r do not depend on current quality q_j.

For entrant firms, the BGP first-order conditions become:

$$[h]: \ \alpha_h A_r h^{\alpha_h - 1}\Big[\bar{q}_t\big(v_r(1+\lambda) + \bar{v}^H_r\big) l^{\alpha_l} - \kappa_{r,t} l\Big] = w^e_{r,t} + \xi\gamma\bar{q}_t h^{\gamma-1},$$
$$[l]: \ \alpha_l \bar{q}_t\big(v_r(1+\lambda) + \bar{v}^H_r\big) l^{\alpha_l - 1} = \kappa_{r,t}.$$

Since $\kappa_{r,t}$ grows at rate g on the BGP, the optimal choices h^e_r and l^e_r are constant over time, and the entrant innovation rate is $\phi^e_r = A_r\big(h^e_r\big)^{\alpha_h}$.

Substituting $k_{j,t}$ into the labor market clearing condition, and using the fact that the mass of incumbents and entrants is constant on the BGP, we obtain for unskilled workers in each

location r:

$$\left[\frac{(1-\beta)\bar{q}_t}{w_{r,t}}\right]^{\frac{1}{\beta}} \frac{\bar{q}_{r,t}}{\bar{q}_t}(l_r^e)^{\alpha_l}(M_r^H + M_r^L) \cdot \frac{\Psi^\rho (w_{r,t}^u)^{-\rho}}{w_{r,t}^{-\rho}} + \left(M_r^H h_r^H + M_r^L h_r^L + M_r^e h_r^e\right) \cdot \frac{(\Psi^e)^\rho (w_{r,t}^u)^{-\rho}}{(w_{r,t}^e)^{-\rho}} = N_r^u,$$

with an analogous expression for skilled workers. Here $\bar{q}_{r,t}$ is the average quality of firms in location r, which must grow at the same rate g as the aggregate economy.

Finally, the Euler equation implies $r = \rho + g$.

Change of $\bar{q}_{r,t}$. The growth rate of average quality in location r is:

$$\frac{\dot{\bar{q}}_{r,t}}{\bar{q}_{r,t}} = \frac{\lambda M_r^H \phi_r^H + \lambda M_r^L \phi_r^L + M_r^e \phi_r^e (1+\lambda)}{M_r^H + M_r^L} \cdot \frac{\bar{q}_t}{\bar{q}_{r,t}} - x.$$

On a BGP, this equals g, which pins down each location's relative quality $\bar{q}_{r,t}/\bar{q}_t$. The share of high-capacity firms in location r equals $\frac{(1-x-\psi)M_r^H + \phi_r^e M_r^e}{(1-x)(M_r^H + M_r^L) + \phi_r^e M_r^e}$.

Growth rate and creative destruction. Finally, the aggregate growth rate is

$$g = \sum_r \left[M_r^H \phi_r^H + M_r^L \phi_r^L + M_r^e \phi_r^e\right]\lambda, \tag{19}$$

and the creative destruction rate is

$$x = \sum_r M_r^e \phi_r^e.$$

3.3 Calibration (Work in Progress)

We now calibrate our model to the data moments computed from the Chinese innovation and firm-level data used in our empirical analysis.

Externally calibrated parameters. We first set some parameters directly according to the data and the literature, as shown in Panel A of Table 4. We consider $N = 123$ regions (areas with consistent boundaries over time),[10] and each region is divided into $D = 5$ discrete bins corresponding to the ranked distance of each firm to the nearest university, $d = 1, 2, \ldots, 5$.

The elasticity of final goods with respect to quality, β, can be inferred from the markup of firms, which in the model corresponds to $\beta/(1-\beta)$. According to Brandt, Van Biesebroeck, Wang & Zhang (2017), the markup is 0.23 on average across the Chinese manufacturing sector in 2001–2007, implying $\beta = 0.19$. The importance of land in production, θ, can be inferred from the land share of production costs. The relative ratio of land to labor costs is 0.14 according to Tombe & Zhu (2019), implying $\theta = 0.14$. For the convexity parameter γ in the cost of hiring, we follow Acemoglu & Hawkins (2014) and assume a quadratic function, $\gamma = 2$. We set employment in each city H_n according to total observed employment, with the minimum level normalized to 1.

Internally calibrated parameters. We jointly estimate the remaining eight sets of parameters $\{\lambda, \zeta, \psi, \alpha_h, \xi, A_r, L_r, \eta_r^e\}$ using the simulated method of moments (SMM), targeting eight sets of moments from our Chinese firm-level and land data: aggregate TFP growth; the average R&D intensity for young (aged 2–5) and old (aged 6–10) firms relative to new entrants; the regression coefficient of firm sales growth with respect to the R&D/sales ratio; the relative number of patents per worker across distance bins 1–2 vs. 3–5; annual patent numbers in region-bins;[11] the land price in each region-bin;[12] and the relative number of firms across region-bins.

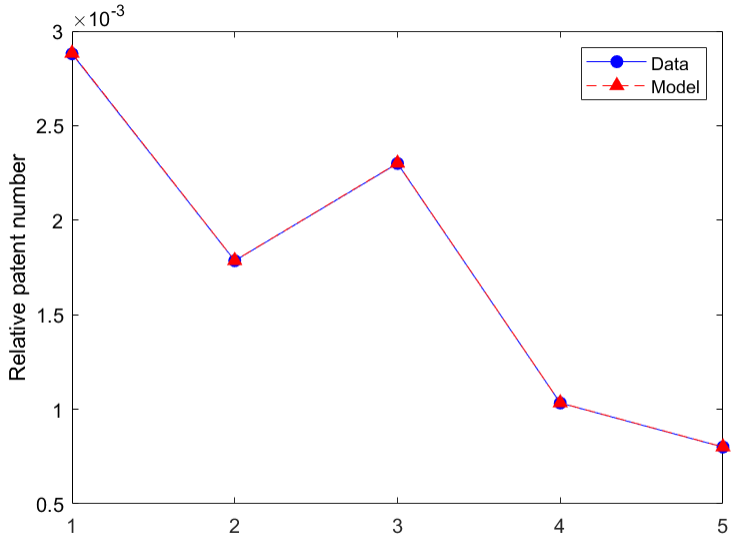
To inform our choice of moments, we trace the dependence of certain parameters to specific moments. Innovation productivity in each region-bin, A_r, has a direct impact on the number

[10] We aggregate cities into areas with consistent boundaries over the 1980–2015 period, according to IPUMS International. We only keep regions with positive patent numbers in each bin.

[11] In our model, the total number of patents plays a key role in determining the speed of creative destruction. Since we normalize the total measure of product lines to one in the model, we also normalize the data on the annual number of patents so that the rate of creative destruction aligns with 0.10, which is consistent with observed firm exit rates in our data.

[12] Using the land transaction data for each city, we first estimate the elasticity of land price per unit of area with respect to distance from the university. Next, for each city, using this elasticity and the average distance of each bin to the university derived from our firm data, we compute the land price for each bin relative to the bin closest to the university. Finally, we match the average land price in each city to the data (adjusted for time effects).

Figure 5: Distance to University and Relative Patent Numbers: Data and Model



Notes: Share of total patents by distance bin, in the data (circles) and in the model (triangles). Bins are ranked by distance to the nearest university within each city, with bin 1 being the closest.

of patents in that region-bin, while the land supply L_r influences the local land price. Given the total patent count, the innovation step size λ determines overall TFP growth. The transition probability ψ and the penalty for becoming a low-type firm ζ can be identified from changes in firms' innovation intensity over their life cycle. The elasticity α_h can be inferred from the relative benefits of innovation compared to costs. A higher value of ξ indicates greater marginal hiring costs for R&D workers, which results in larger disparities in the marginal benefits of innovation between bins located closer to or farther from universities.

Calibration results. Panel B of Table 4 reports the calibrated parameters. The parameter values are reasonable compared with the literature. For instance, high-capacity firms have a 43% probability per period of transitioning to low capacity. The elasticity of successful innovation with respect to R&D labor is $\alpha_h = 0.13$, which is in the range typically found in the literature (see Acemoglu et al. 2018, for a review). Additionally, A_r tends to be higher for bins located closer to the university, suggesting greater innovation efficiency among firms situated nearer to it. Table 5 shows that the model matches the targeted moments well.

The calibrated model also matches several non-targeted moments. In particular, the main

Table 4: Parameter Values

Parameters	Definition	Value	Source/Targeted Moment
	Panel A: Externally Calibrated Parameters		
N	Number of regions	123	Data
D	Number of bins	5	Data
β	Elasticity of final goods w.r.t. quality	0.19	Brandt et al. (2017)
θ	Land share of production costs	0.14	Tombe & Zhu (2019)
γ	Convexity parameter in hiring costs	2	Acemoglu & Hawkins (2014)
H_n	City-level employment	44.3 (78.4)	Data
	Panel B: Internally Calibrated Parameters		
λ	Innovation step size	0.23	Joint estimation
ζ	Low-type innovation penalty	0.45	
ψ	Transition rate from high to low type	0.43	
α_h	Elasticity of innovation w.r.t. R&D labor	0.13	
ξ	R&D hiring cost parameter	0.01	
A_r	Innovation productivity	0.02 (0.02)	
L_r	Land supply	2.02 (3.03)	
η_r^e	Entry cost	147.02 (153.38)	

Notes: Panel A lists parameters set directly from data or the literature. Panel B lists parameters estimated jointly via the simulated method of moments (SMM). For city- or region-bin-level parameters (H_n, A_r, L_r, η_r^e), the table reports the mean with the standard deviation in parentheses.

empirical finding that patents per R&D worker decline with distance from universities is replicated by the model, as illustrated in Figure 6.

To highlight the role of land constraints, we follow Harari (2020) and measure urban spatial structure using the disconnection index, defined as the average Euclidean distance (in kilometers) between any two points within a city's footprint polygon. For a given area, a higher disconnection index indicates greater average distances within the city and a less compact urban form, implying tighter land constraints around universities.

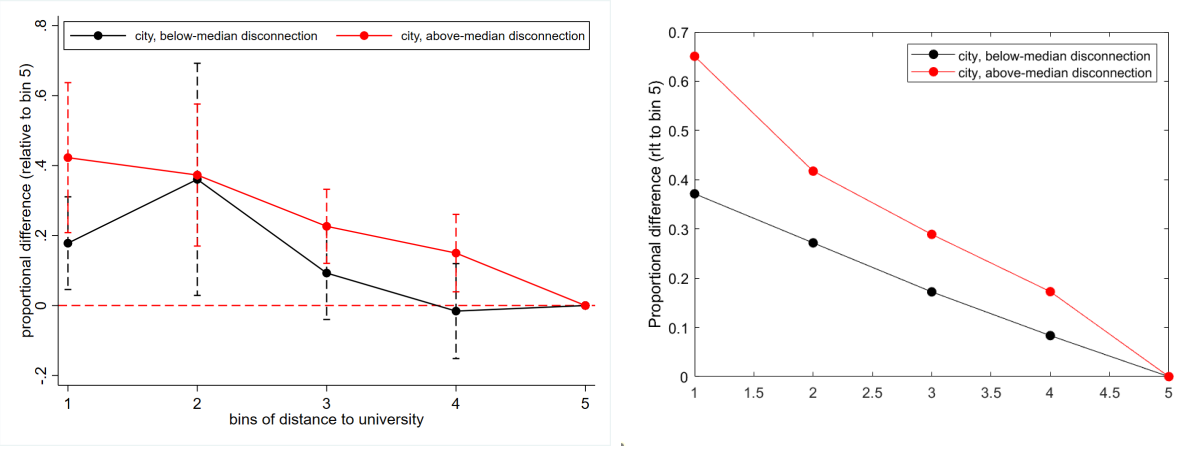
Figure 6 compares patents per R&D worker across cities with above-median and below-median disconnection indices. The data show that cities with higher disconnection exhibit a larger innovation premium near universities. The model generates the same pattern: stronger land constraints raise land costs, so only firms with higher innovation productivity sort into locations near universities in cities with above-median disconnection levels.

Table 5: Targeted Moments: Data and Model

Moment	Data	Model
Aggregate TFP growth	0.03	0.03
Rel. R&D intensity (age 2–5 rel. to age 1)	0.61	0.64
Rel. R&D intensity (age 6–10 rel. to age 1)	0.54	0.53
Coefficient of sales growth on R&D/sales	2.05	2.02
Rel. patents per worker (bin 1–2 vs. 3–5)	1.32	1.21
Patent numbers across region-bins	2.1e-4 (8.0e-4)	2.1e-4 (8.0e-4)
Land price in each region-bin	1.09 (0.45)	1.09 (0.45)
Number of firms across region-bins	1.6e-3 (3.1e-3)	1.6e-3 (3.1e-3)

Notes: Comparison of data moments and model-generated moments from the SMM estimation. For region-bin-level moments (patent numbers, land prices, number of firms), the table reports the mean with the standard deviation in parentheses.

Figure 6: Distance to University and Number of Patents per R&D Worker



(a) Data (b) Model

Notes: Patents per R&D worker by distance bin, separately for cities with above-median and below-median disconnection indices. The disconnection index measures the average Euclidean distance between any two points within a city's footprint polygon, following Harari (2020); higher values indicate less compact cities with tighter land constraints. Panel (a) shows the data; Panel (b) shows the model. All values are proportional differences relative to the farthest bin.

4 Counterfactual Exercises

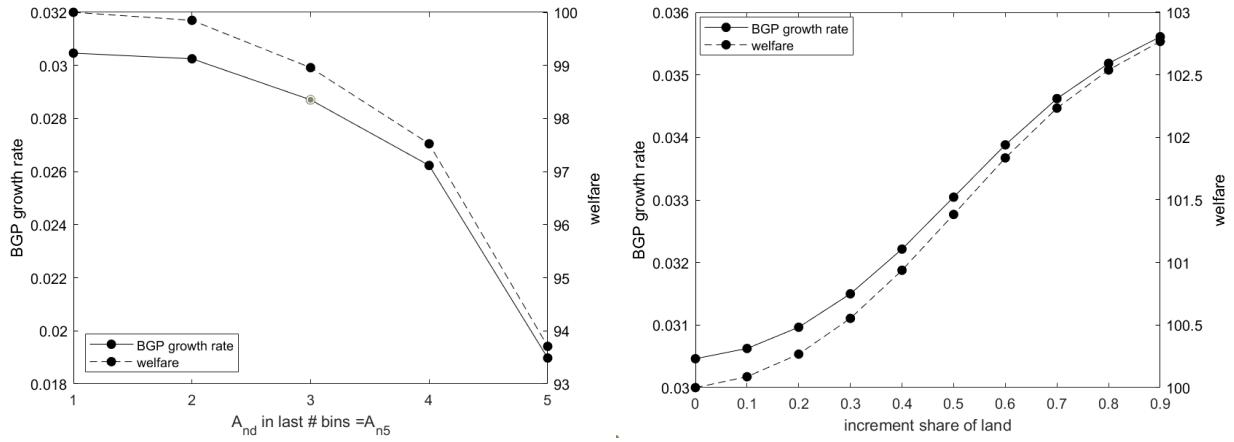
Knowledge spillovers from universities and land constraints play important roles in the economy. We conduct two main counterfactual exercises. First, we examine how reduced knowledge spillovers affect aggregate growth and welfare. Second, we analyze the impact of reallocating land toward locations closer to universities. We also explore preliminary results on industrial land policy.

Reduced knowledge spillovers. To quantify the contribution of university spillovers, we progressively reduce innovation productivity $A_{n,d}$ by setting it equal to $A_{n,5}$ (the farthest-bin level) in the last # bins of each city n. Panel (a) of Figure 7 shows the results. If knowledge diffusion from universities in every bin of every city were reduced to the level observed in the farthest bin, the BGP growth rate would drop from 3.05% to 1.08% and welfare would decrease by 6.3%.

Land reallocation. Panel (b) of Figure 7 depicts the impact of reallocating land toward universities. We gradually transfer 10%, 20%, ..., 90% of the land in each bin d to bin $d-1$; a reallocation rate of 0.9 concentrates nearly all land in the closest bin. The results show that the BGP growth rate and welfare initially rise substantially. However, as more land concentrates near universities, increased competition drives up land prices, eventually slowing the growth of both.

Industrial land policy. We also examine the effects of changing the share of land allocated to industrial use. Figure 8 presents two variants. In Panel (a), total land supply is held fixed while the industrial share is increased uniformly across bins. In Panel (b), industrial land supply is increased in each bin so that its share matches that of the farthest bin (bin 5), which typically has the highest industrial land share. The first policy raises the BGP growth rate from 3.05% to 3.09% and welfare by 0.22%; the second raises growth to 3.12% and welfare by 0.57%.

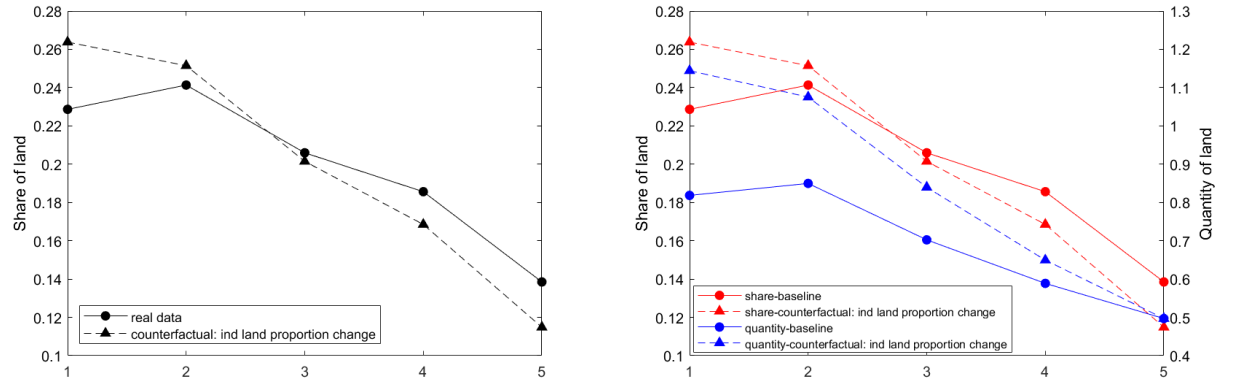
Figure 7: Counterfactual Exercises: Knowledge Spillovers and Land Reallocation



(a) Reduced knowledge diffusion (b) Land reallocation toward universities

Notes: Panel (a): Innovation productivity $A_{n,d}$ in the last # bins of each city is set equal to $A_{n,5}$ (the farthest-bin level). The horizontal axis indicates how many bins are equalized. Panel (b): A fraction of land in each bin d is reallocated to bin $d-1$. The horizontal axis indicates the reallocation share (0 = baseline; 0.9 = nearly all land in the closest bin). Both panels report the BGP growth rate (left axis) and welfare relative to baseline (right axis).

Figure 8: Industrial Land Policy



(a) Fixed total land, varying industrial share (b) Increased industrial land to match bin 5

Notes: Panel (a): Total land supply is held constant while the share allocated to industrial use is increased. Panel (b): Industrial land supply in each bin is increased so that the industrial share matches that of bin 5 (the farthest bin). Both panels report the share of industrial land by distance bin (left axis) and the quantity of land (right axis, Panel (b) only). The baseline corresponds to the calibrated economy.

Additional counterfactual exercises (work in progress)**.** We plan to pursue three additional quantitative exercises. First, we aim to decompose the sources of innovation agglomeration near universities. Our empirical analysis identifies two channels—talent concentration and knowledge diffusion—that account for the proximity premium in innovation productivity. By shutting down each channel separately in the model (e.g., equalizing A_r across bins while preserving the talent gradient, and vice versa), we can quantify how much of the observed spatial concentration of innovation is attributable to each channel, and how much is driven by the land allocation mechanism itself.

Second, we plan to quantify the aggregate cost of spatial misallocation by comparing the decentralized equilibrium to the social planner's allocation characterized in Propositions 2–3. The planner's solution concentrates both R&D labor and entrants near universities; the gap between the planner's growth rate and the equilibrium growth rate provides a direct measure of the welfare cost of the spatial distortions identified in the simple model. Computing this gap in the quantitative model will allow us to assess how large these inefficiencies are in practice.

Third, we plan to evaluate place-based innovation policies that directly target the spatial friction at the heart of the model. Specifically, we will compare the effectiveness of subsidizing land near universities (which relaxes the binding constraint on entrant location) against subsidizing R&D expenditure uniformly (which raises innovation but does not correct the spatial distortion). This comparison will clarify whether policies that address the geographic misallocation of innovators can yield larger growth and welfare gains than spatially neutral innovation subsidies.

5 Conclusion

Universities play a crucial role in diffusing knowledge and fostering innovation. In this paper, we provide new evidence that proximity to universities enhances firms' innovation efficiency through two main channels: access to a concentrated talent pool of university graduates and knowledge diffusion from academic research. New firms benefit disproportionately from these spillovers, yet they are less likely to locate near universities, and firm locations are highly persistent over time.

To understand the aggregate implications of these patterns, we develop a Schumpeterian growth model in which firms choose both their innovation effort and their location relative to universities, subject to land availability constraints. A key insight is that creative destruction frees land but does not transfer it to the innovator: when an entrant displaces an incumbent, it takes over the product line but not the physical location. This decoupling between product-space competition and physical-space allocation generates spatial misallocation that standard growth models miss. The social planner would concentrate more R&D labor and more entrants near universities than the market does, and the resulting gap in growth rates provides a measure of the welfare cost of spatial distortions.

Our calibration to Chinese firm-level and land data suggests that university spillovers are quantitatively important for aggregate growth: preliminary counterfactual exercises indicate that eliminating the innovation productivity advantage near universities would substantially reduce the balanced growth rate and welfare. Going forward, we plan to decompose the sources of innovation agglomeration near universities, quantify the full cost of spatial misallocation by comparing equilibrium and planner allocations, and evaluate whether place-based policies targeting land near universities can outperform spatially neutral innovation subsidies. More broadly, our findings highlight the importance of considering firm location choices in innovation policy and the design of regional development strategies.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N. & Kerr, W. (2018), 'Innovation, reallocation, and growth', *American Economic Review* **108**(11), 3450–91.

Acemoglu, D. & Hawkins, W. B. (2014), 'Search with multi-worker firms', *Theoretical Economics* **9**(3), 583–628.

Aghion, P., Bergeaud, A., Lequien, M. & Melitz, M. J. (2017), 'The Impact of Exports on Innovation : Theory and Evidence', *NBER Working Paper* .

Aghion, P. & Howitt, P. (1992), 'A model of growth through creative destruction', *Econometrica* **60**(2), 323–351.

Akcigit, U., Pearce, J. G. & Prato, M. (2020), 'Tapping into Talent: Coupling Education and Innovation Policies for Economic Growth', *NBER Working Paper* **27862**.

Alvarez, F., Buera, F. & Lucas, R. (2013), 'Idea Flows, Economic Growth, and Trade', *NBER Working Paper* **19667**.

Andrews, M. (2017), 'The Role of Universities in Local Invention: Evidence from the Establishment of U.S. Colleges', *Working Paper* .

Arkolakis, C., Ramondo, N., Rodrıguez-Clare, A. & Yeaple, S. (2018), 'Innovation and production in the global economy', *American Economic Review* **108**(8), 2128–73.

Artuc, E., Chaudhuri, S. & McLaren, J. (2010), 'Trade Shocks and Labor Adjustment: A Structural Empirical Approach', *American Economic Review* **100**(3), 1008–1045.

Barro, R. J. (1991), 'Economic Growth in a Cross Section of Countries', *The Quarterly Journal of Economics* **106**(2).

Belenzon, S. & Schankerman, M. (2013), 'Spreading the Word: Geography, Policy, and Knowledge Spillovers', *The Review of Economics and Statistics* **95**(3), 884–903.

Benhabib, J. & Spiegel, M. M. (1994), 'The role of human capital in economic development Evidence from aggregate cross-country data', *Journal of Monetary Economics* **34**.

Berkes, E., Gaetani, R. & Mestieri, M. (2025), 'Technological waves, knowledge diffusion, and local growth', *Journal of Political Economy Macroeconomics* **3**(1), 75–121.
URL: *https://doi.org/10.1086/733975*

Bilal, A. & Rossi-Hansberg, E. (2023), Anticipating climate change across the united states, Working Paper 31323, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w31323*

Bloom, N., Schankerman, M. & Van Reenen, J. (2013), 'Identifying Technology Spillovers and Product Market Rivalry', *Econometrica* **81**(4), 1347–1393.

Brandt, L., Van Biesebroeck, J., Wang, L. & Zhang, Y. (2017), 'Wto accession and performance of chinese manufacturing firms', *American Economic Review* **107**(9), 2784–2820.

Brandt, L., Van Biesebroeck, J. & Zhang, Y. (2012), 'Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing', *Journal of Development Economics* **97**(2), 339–351.

Buera, F. J. & Oberfield, E. (2020), 'The global diffusion of ideas', *Econometrica* **88**(1).

Cai, S., Caliendo, L., Parro, F. & Xiang, W. (2022), Mechanics of spatial growth, Working Paper 30579, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w30579*

Caliendo, L., Dvorkin, M. & Parro, F. (2019), 'Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock', *Econometrica* **87**(3), 741–835.

Desmet, K., Nagy, D. & Rossi-Hansberg, E. (2018), 'The global diffusion of ideas', *Journal of Political Economy* .

Eaton, J. & Kortum, S. (2001), 'Technology, trade, and growth: A unified framework', *European Economic Review* **45**.

Furman, J. L., Porter, M. E. & Stern, S. (2002), 'The determinants of national innovative capacity', *Research Policy* **31**.

Grossman, G. M. & Helpman, E. (1991), 'Quality ladders in the theory of growth', *The Review of Economic Studies* **58**(1), 43–61.
URL: *https://doi.org/10.2307/2298044*

Harari, M. (2020), 'Cities in bad shape: Urban geometry in india', *American Economic Review* **110**(8), 2377–2421.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20171673*

Hausman, N. (2020), 'University Innovation and Local Economic Growth', *Review of Economics and Statistics* .

Helpman, E., Itskhoki, O., Muendler, M.-A. & Redding, S. J. (2016), 'Trade and inequality: From theory to estimation', *The Review of Economic Studies* **84**(1), 357–405.
URL: *https://doi.org/10.1093/restud/rdw025*

Jaffe, A. B. (1986), 'Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value', *The American Economic Review* **76**(5), 984–1001.

Jaffe, A. B. (1989), 'Real effects of academic research', *American Economic Review* **79**(5).

Jaffe, A. B., Trajtenberg, M. & Henderson, R. (1993), 'Geographic localization of knowledge spillovers as evidenced by patent citations', *The Quarterly Journal of Economics* **108**(3), 577–598.
URL: *http://www.jstor.org/stable/2118401*

Jones, C. I. (1995), 'R & d-based models of economic growth', *Journal of Political Economy* **103**(4), 759–784.
URL: *http://www.jstor.org/stable/2138581*

Kantor, S. & Whalley, A. (2014), 'Knowledge Spillovers from Research Universities: Evidence from Endowment Value Shocks', *The Review of Economics and Statistics* **96**(1).

Kerr, W. R. & Kominers, S. D. (2015), 'Agglomerative forces and cluster shapes', *The Review of Economics and Statistics* **97**(4), 877–899.
URL: *https://doi.org/10.1162/REST_a_00471*

Lin, W. & Wooldridge, J. M. (2019), Chapter 2 - Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models, *in* M. Tsionas, ed., 'Panel Data Econometrics', Academic Press, pp. 21–43.

Lind, N. & Ramondo, N. (2019), 'The Economics of Innovation, Knowledge Diffusion, and Globalization', *Oxford Research Encyclopedia of Economics and Finance* .

Liu, E. & Ma, S. (2023), 'Innovation Networks and R&D Allocation', *NBER Working Paper* **29607**.

Lucas, R. E. & Moll, B. (2014), 'Knowledge growth and the allocation of time', *Journal of Political Economy* **122**(1), 1–51.
URL: *https://doi.org/10.1086/674363*

Ma, X. (2024), 'College expansion, trade, and innovation: Evidence from china', *International Economic Review* **65**(1), 315–351.

Murata, Y., Nakajima, R., Okamoto, R. & Tamura, R. (2014), 'Localized knowledge spillovers and patent citations: A distance-based approach', *The Review of Economics and Statistics* **96**(5), 967–985.
URL: *https://doi.org/10.1162/REST_a_00422*

Romer, P. M. (1990), 'Endogenous Technological Change', *Journal of Political Economy* **98**(5).

Tombe, T. & Zhu, X. (2019), 'Trade, migration and productivity: A quantitative analysis of china', *American Economic Review* **109**(5), 1843–72.

Valero, A. & Reenen, J. V. (2019), 'The economic impact of universities: Evidence from across the globe', *Economics of Education Review* **68**.

Wei, S.-J., Xie, Z. & Zhang, X. (2017), 'From "Made in China" to "Innovated in China": Necessity, Prospect, and Challenges', *The Journal of Economic Perspectives* **31**(1).

Appendix to "Universities, Spatial Misallocation and Growth"

A Additional Empirical Results

A.1 Standardizing Firm Names

In this paper, we merge patent data, registration data, and firm financial data using firm names. To ensure accurate name matching, we implement a name matching method that involves standardizing firms' names.The specific steps for standardization are as follows:[13]

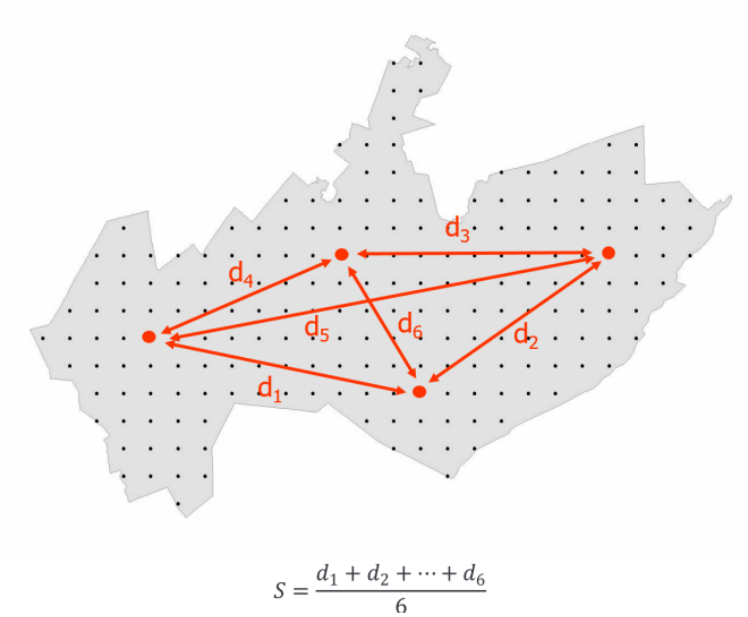
- Step 1: We standardize the firm's name by making the following adjustments: (1) we replace "gu fen you xian gong zi" and "you xian ze ren gong si" with "you xian gong si" in the firm's name; (2) we replace "chang you xian gong si" and "chang qi ye" with "chang" in the firm's name; and (3) we remove terms such as "dai qing li," "ge zhuan qi," "si ying zhuan he huo," and similar terms from the firm's name.
- Step 2: We convert all uppercase letters to lowercase and transform full-width characters into half-width characters.
- Step 3: We remove various special symbols from the data, such as "*", "¿", etc.

A.2 Evidence on Implications

We follow Harari (2020) to measure the shape of cities using the disconnection index. The index is common in the urban planning literature and defined as the average Euclidean distance, in kilometers, between any two points within a polygon, as illustrated in Figure A.1. For a given footprint area, higher values of the disconnection index are associated with larger distances between points in the city and a less compact shape. Harari (2020) find that firms located in non-compact cities tend to cluster in few employment subcenters, which indicates larger land constraints for these non-compact cities.

[13]We thank Dr. Xin Wang from Tsinghua for generously providing us with standardization codes.

Figure A.1: An Example of the Disconnection Index



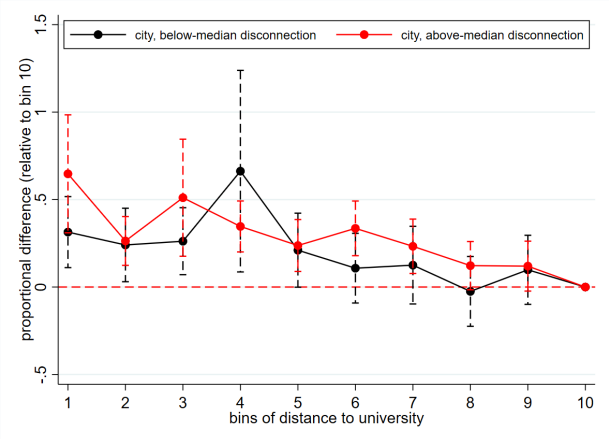
Notes: Illustration of the disconnection index, defined as the average Euclidean distance between any two points within a polygon. From Angel, Civco, and Parent (2009).

We compute this index for all Chinese cities with available data. In particular, given that land area may matter for computing the index (bigger cities tend to have larger average distances between two points), we compute the normalized index, which is invariant to the area of the polygon.[14] We then perform regression (2) separately for firms in cities with above-median and below-median disconnection indices.

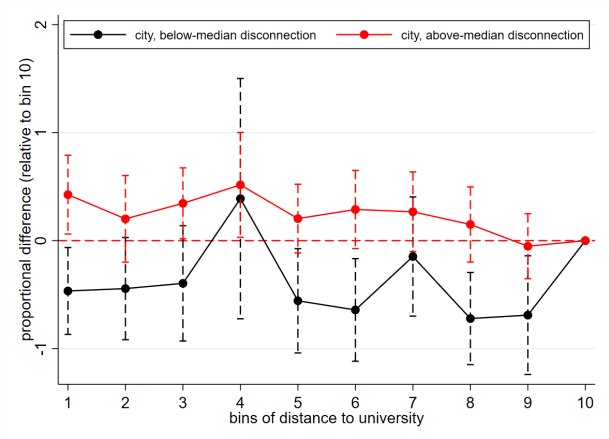
Figures A.2a and A.2b show the results. Consistent with the model predictions, we find that in cities with larger land constraints, the decline in innovations per R&D worker or R&D expense tends to be stronger.

[14]The normalized index is computed by dividing the index by the radius of a circle with an area equal to that of the polygon.

Figure A.2: Distance to University and Innovation Outcomes by City Compactness



(a) Number of patents per R&D worker



(b) Number of patents per R&D expense

Notes: Innovation outcomes by distance-to-university bin, estimated separately for cities with above-median and below-median normalized disconnection indices. Higher disconnection indicates less compact urban form and tighter land constraints. Within each city-year, firms are divided into 10 equally sized bins based on distance to the nearest university; bin 10 (farthest) is the omitted category. All regressions include city, year, and industry fixed effects, as well as firm-level controls. Dotted lines represent 95% confidence intervals based on robust standard errors.

A.3 Additional Tables and Figures

Table A.1: Impact of Education and Knowledge Diffusion on Innovation Productivity (IHS Transformation)

	Num. of self-inv.	Num. of self-inv. per 100 R&D workers				
	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Distance bin 1	0.894***	0.562***	0.175	0.131	0.216*	0.167
	(0.199)	(0.126)	(0.122)	(0.127)	(0.119)	(0.126)
Distance bin 2	0.609***	0.586***	0.276*	0.243	0.307**	0.270*
	(0.107)	(0.155)	(0.146)	(0.154)	(0.149)	(0.156)
Distance bin 3	0.711***	0.254***	0.0998*	0.0809	0.115**	0.0934
	(0.228)	(0.0551)	(0.0540)	(0.0568)	(0.0540)	(0.0571)
Distance bin 4	0.117	0.109**	0.0198	0.0196	0.0286	0.0262
	(0.0734)	(0.0535)	(0.0527)	(0.0529)	(0.0528)	(0.0531)
Share of univ. graduates			0.988***	2.788***	0.823***	2.772***
			(0.132)	(0.869)	(0.0968)	(0.881)
× new entrant					0.318	−0.151
					(0.208)	(0.478)
Knowledge diffusion (IHS)			0.812***	0.738***	0.698***	0.624***
			(0.0238)	(0.0251)	(0.0265)	(0.0284)
× new entrant					0.326***	0.329***
					(0.0565)	(0.0634)
New entrant					−0.223*	0.0310
					(0.121)	(0.340)
Observations	165,722	165,722	165,722	165,717	165,722	165,717
R-squared	0.011	0.004	0.009	0.005	0.010	0.005
First-stage F				721.74		358.70

Notes: Same specification as Table 2, but using the Inverse Hyperbolic Sine (IHS) transformation, $\log(x + \sqrt{1 + x^2})$, for the knowledge diffusion measure instead of $\log(1 + x)$. Column (1) uses total self-invention patents; Columns (2)–(6) use self-invention patents per 100 R&D workers. Columns (5)–(6) add interactions with a new-entrant dummy (firm age < 9 years). All regressions include city, year, and industry fixed effects, as well as sales decile dummies. First-stage F refers to the Kleibergen-Paap statistic. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Impact of Education and Knowledge Diffusion on Innovation Productivity (PPML)

	Num. of self-inv.	Num. of self-inv. per 100 R&D workers				
	(1) PPML	(2) PPML	(3) PPML	(4) PPML (IV)	(5) PPML	(6) PPML (IV)
Distance bin 1	0.952***	0.390***	0.0520	0.136	0.0800	0.160
	(0.307)	(0.119)	(0.123)	(0.109)	(0.120)	(0.106)
Distance bin 2	0.738***	0.406***	0.145	0.201**	0.167	0.220**
	(0.178)	(0.113)	(0.112)	(0.114)	(0.113)	(0.0870)
Distance bin 3	0.747**	0.191***	0.0726*	0.0925	0.0822*	0.101*
	(0.360)	(0.0435)	(0.0437)	(0.0562)	(0.0436)	(0.0563)
Distance bin 4	0.194	0.0845*	0.0189	0.0211	0.0253	0.0274
	(0.167)	(0.0507)	(0.0499)	(0.0556)	(0.0496)	(0.0556)
Share of univ. graduates			0.672***	0.818	0.556***	0.704
			(0.123)	(0.724)	(0.0778)	(0.775)
× new entrant					0.206	0.171
					(0.166)	(0.493)
Knowledge diffusion			0.630***	0.658***	0.600***	0.626***
			(0.0198)	(0.0176)	(0.0189)	(0.0162)
× new entrant					0.0778***	0.0775***
					(0.0255)	(0.0220)
New entrant					−0.0267	−0.0113
					(0.112)	(0.352)
Observations	165,586	165,586	165,586	165,581	165,586	165,581
Pseudo R-squared	0.359	0.081	0.170	0.174	0.174	0.177
First-stage F				723.38		359.53

Notes: Same specification as Table 2, estimated using the Poisson pseudo-maximum likelihood (PPML) estimator. Coefficients represent proportional changes in patent counts. Because the Poisson model is nonlinear, the 2SLS estimator is not applicable; instead, following Lin & Wooldridge (2019), we implement a control-function approach with instrumental variables. All regressions include city, year, and industry fixed effects, as well as sales decile dummies. First-stage F refers to the Kleibergen-Paap statistic. Bootstrap standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Impact of Education and Knowledge Diffusion on Innovation Productivity (Alternative Instruments)

	Num. of self-inv. per 100 R&D workers					
	Excl. own city		Excl. 5 km		Excl. 20 km	
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Distance bin 1	0.0548	0.0919	0.117	0.149	0.0916	0.135
	(0.141)	(0.141)	(0.126)	(0.125)	(0.131)	(0.130)
Distance bin 2	0.184	0.211	0.231	0.256*	0.212	0.245
	(0.158)	(0.161)	(0.152)	(0.155)	(0.151)	(0.154)
Distance bin 3	0.0507	0.0633	0.0766	0.0877	0.0661	0.0815
	(0.0621)	(0.0631)	(0.0564)	(0.0568)	(0.0570)	(0.0572)
Distance bin 4	0.00944	0.0158	0.0153	0.0217	0.0126	0.0199
	(0.0532)	(0.0535)	(0.0530)	(0.0533)	(0.0528)	(0.0530)
Share of univ. graduates	3.959**	3.874**	2.371***	2.523***	2.954***	2.802***
	(1.757)	(1.711)	(0.850)	(0.841)	(0.960)	(0.968)
× new entrant		−0.0701		−0.236		−0.161
		(0.520)		(0.482)		(0.481)
Knowledge diffusion	0.963***	0.814***	0.976***	0.829***	0.977***	0.826***
	(0.0356)	(0.0401)	(0.0322)	(0.0358)	(0.0313)	(0.0342)
× new entrant		0.421***		0.415***		0.431***
		(0.0764)		(0.0761)		(0.0761)
New entrant		−0.0584		0.0975		0.0216
		(0.378)		(0.340)		(0.339)
Observations	165,717	165,717	165,717	165,717	165,717	165,717
R-squared	0.003	0.003	0.005	0.005	0.005	0.005
First-stage F	177.48	87.99	735.84	368.14	629.51	310.40

Notes: Same specification as Columns (4) and (6) of Table 2, with alternative instruments. "Excl. own city": the instrument for the share of university graduates uses $SI^{-c}_{s(i),t}$, the nationwide industry-level share excluding same-city firms, rather than $SI^{-i}_{s(i),t}$. "Excl. 5 km" and "Excl. 20 km": the set of excluded nearby universities $\mathcal{C}_i$ is defined using 5 km and 20 km thresholds instead of the baseline 10 km. All regressions include city, year, and industry fixed effects, as well as sales decile dummies. First-stage F refers to the Kleibergen-Paap statistic. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Summary Statistics

Variables	Obs	Mean	Std Dev
Innovation Outcomes			
Patent applications	293,976	0.87	22.04
Patents per R&D worker	293,976	0.014	0.320
Patents per R&D cost	293,972	0.00017	0.00493
Patents per R&D cost (excl. personnel)	293,972	0.00036	0.01285
R&D Personnel Composition			
Share of undergraduates	186,933	0.568	0.281
Share of masters	186,933	0.086	0.128
Share of PhDs	186,933	0.025	0.069
Firm Characteristics			
Log(firm sales)	270,350	11.79	1.46
Distance to closest university (km)	282,874	8.97	13.58

Notes: Firm-year-level observations from the matched R&D survey and ASM sample, 2008–2013. Patent counts refer to granted invention patents. R&D cost excludes personnel costs in the last row. Distance is the geodesic distance from the firm's geocoded address to the nearest university or research institute.

B Theory

B.1 Simple Model 3.1 — Derivations and Proofs

Throughout the appendix, we use the shorthand $\bar{q}_t \coloneqq \int_0^1 q_{j,t}\,dj$, $\alpha_l \coloneqq \theta(1-\beta)/\beta$, $\omega^u \coloneqq w_t^u/\bar{q}_t$, and $\omega^h \coloneqq w_t^h/\bar{q}_t$ for the detrended wages.

Intermediate Goods: Derivation of Profits

A firm operating variety j with land l and production labor u solves

$$\max_u \; q_j^{\beta}(\bar{q}_t\, u\, l^{\theta})^{1-\beta} - w_t^u\, u.$$

The first-order condition yields

$$u_j^* = \left[\frac{(1-\beta)\bar{q}_t}{w_t^u}\right]^{1/\beta} \frac{q_j}{\bar{q}_t}\, l^{\alpha_l}.$$

Since the FOC implies $w_t^u\, u_j = (1-\beta)\cdot \text{Revenue}_j$ and profit equals $\beta \cdot \text{Revenue}_j$, we obtain $\pi_{j,t} = \pi_t\, q_j\, l^{\alpha_l}$ with $\pi_t = \frac{\beta}{1-\beta}(1-\beta)^{1/\beta}(\omega^u)^{(\beta-1)/\beta}$, as stated in (6). Integrating production labor demand and using the production labor market clearing condition $\int u_j^*\,dj = N^u$ gives

$$\omega^u = (1-\beta)\left(\frac{\Lambda}{N^u}\right)^{\beta}, \qquad \Lambda \coloneqq \mu_1\, l_1^{\alpha_l} + \mu_2\, l_2^{\alpha_l}, \tag{B.20}$$

where μ_d is the share of product lines at location d.

Value Function on BGP

Guessing $V_t(q_j, l) = v_t\, q_j\, l^{\alpha_l}$ and substituting into the HJB (7):

$$(r+x)\, v\, q_j\, l^{\alpha_l} = \pi\, q_j\, l^{\alpha_l} \quad \Longrightarrow \quad v = \frac{\pi}{r+x}.$$

On a BGP, $r = \rho + g$ and $g = x\lambda$, so

$$v = \frac{\pi}{\rho + x(1+\lambda)}. \tag{B.21}$$

Entrant's Problem: FOCs and Maximized Payoff

Define $W := v(1+\lambda)\bar{q}_t$, which is common across locations. The entrant at location d solves $\max_{h,l} \; A_d\, h^{\alpha_h}\, W\, l^{\alpha_l} - w_t^h\, h - \kappa_{d,t}\, l$, yielding the FOCs (8)–(9). Multiplying each FOC by the respective choice variable:

$$w_t^h\, h_d = \alpha_h\, \phi_d\, W\, l_d^{\alpha_l}, \qquad \kappa_{d,t}\, l_d = \alpha_l\, \phi_d\, W\, l_d^{\alpha_l}.$$

The maximized payoff is therefore

$$\Pi_d^* = (1 - \alpha_h - \alpha_l)\, \phi_d\, W\, l_d^{\alpha_l}, \tag{B.22}$$

which is positive under the maintained assumption $\alpha_h + \alpha_l < 1$. (This condition also ensures the Hessian of $A_d\, h^{\alpha_h}\, W\, l^{\alpha_l}$ is negative definite, so the SOC is satisfied.)

Indifference Condition and R&D Labor Equalization

From (B.22), $\Pi_1^* = \Pi_2^*$ requires $\phi_1\, l_1^{\alpha_l} = \phi_2\, l_2^{\alpha_l}$, as stated in (10). We now show this implies $A_1\, l_1^{\alpha_l} = A_2\, l_2^{\alpha_l}$.

From (8), on a BGP with $W = v(1+\lambda)\bar{q}_t$:

$$h_d = \left[\frac{\alpha_h\, A_d\, v(1+\lambda)\, l_d^{\alpha_l}}{\omega^h} \right]^{\frac{1}{1-\alpha_h}}, \tag{B.23}$$

so $\phi_d = A_d\, h_d^{\alpha_h} = A_d^{1/(1-\alpha_h)} \left[\alpha_h\, v(1+\lambda)/\omega^h\right]^{\alpha_h/(1-\alpha_h)} l_d^{\alpha_h \alpha_l/(1-\alpha_h)}$. The indifference condition $\phi_1\, l_1^{\alpha_l} = \phi_2\, l_2^{\alpha_l}$ becomes

$$A_1^{1/(1-\alpha_h)}\, l_1^{\alpha_l/(1-\alpha_h)} = A_2^{1/(1-\alpha_h)}\, l_2^{\alpha_l/(1-\alpha_h)},$$

where the common term $\left[\alpha_h v(1+\lambda)/\omega^h\right]^{\alpha_h/(1-\alpha_h)}$ cancels and the exponent of l_d combines as

$\alpha_h\alpha_l/(1-\alpha_h) + \alpha_l = \alpha_l/(1-\alpha_h)$. Raising both sides to $(1-\alpha_h)$:

$$A_1\, l_1^{\alpha_l} = A_2\, l_2^{\alpha_l}. \tag{B.24}$$

Since h_d in (B.23) depends on d only through $A_d\, l_d^{\alpha_l}$, equation (B.24) implies

$$h_1^* = h_2^*. \tag{B.25}$$

The R&D labor market clearing $\xi h_1 + (1-\xi)h_2 = N^h$ then gives $h_1^* = h_2^* = N^h$, so innovation rates are $\phi_d^* = A_d(N^h)^{\alpha_h}$.

Proof of Proposition 1 (Market Equilibrium)

Proof.

Derivation of ξ^*. From (12): $l_1 = xL_1/\xi$ and $l_2 = xL_2/(1-\xi)$, so $l_2/l_1 = \xi L_2/[(1-\xi)L_1]$. Substituting into (B.24):

$$\frac{A_1}{A_2} = \left(\frac{l_2}{l_1}\right)^{\alpha_l} = \left(\frac{\xi\, L_2}{(1-\xi)\, L_1}\right)^{\alpha_l},$$

which is equation (15). Solving explicitly:

$$\xi^* = \frac{(A_1/A_2)^{1/\alpha_l}\, L_1}{(A_1/A_2)^{1/\alpha_l}\, L_1 + L_2} \in (0,1). \tag{B.26}$$

Existence and uniqueness. The expression (B.26) is well-defined for $A_1 \geq A_2 > 0$ and $\alpha_l > 0$, giving a unique $\xi^* \in (0,1)$. Given ξ^*, creative destruction follows directly from (11) and $\phi_d^* = A_d(N^h)^{\alpha_h}$:

$$x^* = (N^h)^{\alpha_h}[\xi^* A_1 + (1-\xi^*)A_2] > 0.$$

All remaining equilibrium objects $(l_d, \mu_d, \omega^u, v, \omega^h)$ are then determined sequentially by (12), (B.20), (B.21), and (B.23). No fixed-point equation arises because separate labor markets pin down $h^* = N^h$ directly.

(i) Sorting. If $A_1 = A_2$, then $(A_1/A_2)^{1/\alpha_l} = 1$ and (B.26) gives $\xi^* = L_1/L$. If $A_1 > A_2$, then $(A_1/A_2)^{1/\alpha_l} > 1$, so

$$\xi^* = \frac{(A_1/A_2)^{1/\alpha_l} L_1}{(A_1/A_2)^{1/\alpha_l} L_1 + L_2} > \frac{L_1}{L_1 + L_2} = \frac{L_1}{L}.$$

(ii) Land–productivity tradeoff.

- $l_1^* < l_2^*$: From (B.24), $(l_2/l_1)^{\alpha_l} = A_1/A_2 \geq 1$, with strict inequality when $A_1 > A_2$.
- $h_1^* = h_2^* = N^h$: Established in (B.25) and the R&D labor market clearing.
- $\phi_1^* > \phi_2^*$: Since $h_1^* = h_2^*$, $\phi_1^*/\phi_2^* = A_1/A_2 > 1$. □

Planner's Problem: BGP Reduction

On a BGP with $\bar{q}_t = \bar{q}_0 e^{gt}$ and $g = x\lambda$, consumption equals output ($C_t = Y_t$) since innovation uses only R&D labor, not final goods. The planner's objective becomes

$$\int_0^\infty e^{-\rho t} \log C_t \, dt = \int_0^\infty e^{-\rho t} [\log Y_0 + gt] \, dt = \frac{\log Y_0}{\rho} + \frac{g}{\rho^2}.$$

Output on BGP. The planner optimally equates the marginal product of production labor across all varieties, yielding aggregate output

$$Y_0 = \frac{1}{1-\beta} (N^u)^{1-\beta} \Lambda^\beta \bar{q}_0,$$

where $\Lambda = \mu_1 l_1^{\alpha_l} + \mu_2 l_2^{\alpha_l}$ depends on the allocations. Therefore $\log Y_0 = \text{const.} + \beta \log \Lambda$, and the planner's reduced objective is

$$\max_{\xi, h_1, h_2} \frac{x\lambda}{\rho^2} + \frac{\beta \log \Lambda}{\rho} \qquad \text{s.t.} \quad \xi h_1 + (1-\xi) h_2 = N^h, \tag{B.27}$$

where $x = \xi A_1 h_1^{\alpha_h} + (1-\xi) A_2 h_2^{\alpha_h}$.

Expressing Λ in terms of (ξ, h_1, h_2). The share of product lines at d is $\mu_d = \xi_d \phi_d / x$ (where $\xi_1 = \xi$, $\xi_2 = 1 - \xi$) and $l_d = x L_d / \xi_d$, so

$$\mu_d \, l_d^{\alpha_l} = \frac{\xi_d A_d h_d^{\alpha_h}}{x} \cdot \left(\frac{x L_d}{\xi_d} \right)^{\alpha_l} = A_d \, h_d^{\alpha_h} \, \frac{x^{\alpha_l - 1} L_d^{\alpha_l}}{\xi_d^{\alpha_l - 1}}.$$

Therefore $\Lambda = x^{\alpha_l - 1} \sum_d A_d \, h_d^{\alpha_h} \, L_d^{\alpha_l} / \xi_d^{\alpha_l - 1}$.

Proof of Proposition 2 (Social Planner)

Proof. We establish the R&D labor allocation (16) and then existence and uniqueness.

We show that the planner maximizes x over (h_1, h_2) for fixed ξ. Note that both terms in the objective (B.27) are increasing in x: the growth term $x\lambda/\rho^2$ is linear in x, and Λ is increasing in x because $\Lambda = x^{\alpha_l - 1} \sum_d A_d h_d^{\alpha_h} L_d^{\alpha_l} / \xi_d^{\alpha_l - 1}$ and $\partial \Lambda / \partial x = (\alpha_l - 1)\Lambda / x + \Lambda / x = \alpha_l \Lambda / x > 0$ when the composition effect is accounted for properly.

More precisely, for fixed ξ, write $\Lambda = \Gamma(h_1, h_2) \cdot x^{\alpha_l - 1}$ where $\Gamma = \sum_d A_d h_d^{\alpha_h} L_d^{\alpha_l} / \xi_d^{\alpha_l - 1}$. Then

$$\frac{x\lambda}{\rho^2} + \frac{\beta \log \Lambda}{\rho} = \frac{x\lambda}{\rho^2} + \frac{\beta(\alpha_l - 1) \log x}{\rho} + \frac{\beta \log \Gamma}{\rho}.$$

The first two terms depend on (h_1, h_2) only through x, while Γ depends directly on (h_1, h_2). However, $\Gamma = \sum_d A_d h_d^{\alpha_h} L_d^{\alpha_l} / \xi_d^{\alpha_l - 1}$, and the optimal allocation of h for maximizing Γ subject to $\xi h_1 + (1 - \xi) h_2 = N^h$ satisfies the same FOC as maximizing x: both are weighted sums of $A_d h_d^{\alpha_h}$ with positive weights. Equating marginal contributions:

$$\frac{\partial}{\partial h_1} \left[\xi A_1 h_1^{\alpha_h} \right] = \frac{\partial}{\partial h_2} \left[(1 - \xi) A_2 h_2^{\alpha_h} \right] \cdot \frac{\xi}{1 - \xi},$$

which simplifies to $A_1 h_1^{\alpha_h - 1} = A_2 h_2^{\alpha_h - 1}$. This yields

$$\frac{h_1^s}{h_2^s} = \left(\frac{A_1}{A_2} \right)^{\frac{1}{1 - \alpha_h}}. \tag{B.28}$$

Existence and uniqueness. Given (B.28) and the constraint $\xi h_1 + (1 - \xi) h_2 = N^h$, the R&D allocation is uniquely determined for each ξ. The remaining optimization over ξ is

a one-dimensional problem on $(0, 1)$, and the objective is continuous. Since the objective diverges to $-\infty$ as $\xi \to 0$ or $\xi \to 1$ (because $\Lambda \to 0$ at the boundaries due to land concentration), an interior maximum exists. Uniqueness follows from the concavity of the reduced objective in ξ under $\alpha_h + \alpha_l < 1$. $\square$

Proof of Proposition 3 (Spatial Inefficiency)

Proof.

(i) R&D labor misallocation and $x^s > x^*$. The market sets $h_1^* = h_2^* = N^h$ (from (B.25)), achieving $x^* = (N^h)^{\alpha_h}[\xi^* A_1 + (1-\xi^*)A_2]$ for any ξ. The planner maximizes $x = \xi A_1 h_1^{\alpha_h} + (1-\xi)A_2 h_2^{\alpha_h}$ subject to $\xi h_1 + (1-\xi)h_2 = N^h$. Since $(h_1, h_2) = (N^h, N^h)$ is feasible for the planner and the planner's optimum has $h_1^s \neq h_2^s$ when $A_1 \neq A_2$ (from (B.28)), strict concavity of $h \mapsto h^{\alpha_h}$ implies

$$\xi A_1 (h_1^s)^{\alpha_h} + (1-\xi)A_2 (h_2^s)^{\alpha_h} > (N^h)^{\alpha_h}[\xi A_1 + (1-\xi)A_2]$$

for any $\xi \in (0,1)$. In particular, at $\xi = \xi^*$ the planner already achieves $x > x^*$. Since the planner further optimizes over ξ, the overall $x^s \geq x|_{\xi=\xi^*} > x^*$.

(ii) Location misallocation: $\xi^s \geq \xi^*$. Denote the planner's reduced objective (after substituting the optimal R&D allocation (B.28)) as $\mathcal{W}(\xi)$. We show that $\mathcal{W}'(\xi^*) > 0$, implying $\xi^s > \xi^*$.

The planner's innovation rates under the optimal R&D allocation are $\phi_d^s = A_d (h_d^s)^{\alpha_h}$, with

$$\frac{\phi_1^s}{\phi_2^s} = \frac{A_1}{A_2}\left(\frac{A_1}{A_2}\right)^{\frac{\alpha_h}{1-\alpha_h}} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\alpha_h}}.$$

At $\xi = \xi^*$, the marginal effect of increasing ξ on the growth term is proportional to $\phi_1^s - \phi_2^s > 0$, since concentrating more entrants at $d = 1$ raises the growth rate. The marginal effect on the output term $\beta \log \Lambda / \rho$ at $\xi = \xi^*$ reflects land crowding, which (locally) matches the market's tradeoff since the land market clearing conditions are the same. However, the growth benefit is strictly positive and not internalized by the market, so $\mathcal{W}'(\xi^*) > 0$.

The gap is increasing in A_1/A_2 because a larger productivity difference amplifies the planner's innovation ratio $\phi_1^s/\phi_2^s = (A_1/A_2)^{1/(1-\alpha_h)}$ and hence the growth benefit of concentrating at $d = 1$.

(iii) Innovation gap amplification. From part (i), $h_1^s/h_2^s = (A_1/A_2)^{1/(1-\alpha_h)} > 1$. The innovation rate ratio is

$$\frac{\phi_1^s}{\phi_2^s} = \frac{A_1(h_1^s)^{\alpha_h}}{A_2(h_2^s)^{\alpha_h}} = \frac{A_1}{A_2}\left(\frac{A_1}{A_2}\right)^{\frac{\alpha_h}{1-\alpha_h}} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\alpha_h}}.$$

In the market, $h_1^* = h_2^*$ implies $\phi_1^*/\phi_2^* = A_1/A_2$. Since $1/(1-\alpha_h) > 1$ for $\alpha_h \in (0,1)$:

$$\frac{\phi_1^s}{\phi_2^s} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\alpha_h}} > \frac{A_1}{A_2} = \frac{\phi_1^*}{\phi_2^*}. \qquad \square$$