

# The Geography of Innovation in the United States

Weiliang Tan\*  
National Bureau of Economic Research

December 10, 2025

Click [here](#) for the latest version

## Abstract

A defining trend in U.S. innovation is its rising geographic concentration, exemplified by the growth of high-tech clusters such as Silicon Valley. What factors drive this rising spatial concentration, and what are its implications for regional and aggregate growth? Using comprehensive data on patents, firms, and inventors from 1976 to 2018, I find that innovation became more concentrated in high-skill regions only after 1990, with the rise of information and communication technologies (ICT) playing two distinct roles in this process. First, the sudden arrival of influential ICT breakthroughs shifted innovation toward ICT – a sector whose production and innovation are more concentrated in high-skill regions relative to the non-ICT sector. Second, ICT-enabled reductions in communication costs allowed innovation-intensive firms based in high-skill regions to expand geographically, increase their scale, and innovate more. Worker sorting toward high-skill regions in the 1980s facilitated the growth of both the ICT sector and other innovation-intensive firms in these regions. To decompose the contribution of each mechanism to the rising spatial concentration of innovation and its implications for macroeconomic growth, I develop a spatial growth model with endogenous and directed innovation, technology diffusion, and dynamic worker mobility. The model characterizes the *spatial direction of innovation* along the transition path and shows how its steady-state spatial distribution shapes long-run aggregate growth.

**Keywords:** spatial innovation; trade and growth; agglomeration economies; information and communication technologies

---

\*Email: wt289@cornell.edu. I am extremely grateful to my advisors Samuel Kortum, Peter Schott, Scott Stern, Ivan Rudik, and Nancy Chau for valuable advice. I am indebted to Professor Rudik for all his support during my PhD and to Professor Kortum for hosting me at Yale from 2023-2025 and for all his guidance. I thank Andy Bernard, Michael Peters, Costas Arkolakis, Lorenzo Caliendo, Allan Hsiao, Shanjun Li, Ezra Oberfield, Bernardo Ribeiro, Sabrina Peng, and Zebang Xu for helpful conversations.

# 1 Introduction

The central questions in spatial economics are what drives the geographic distribution of economic activity and what are its local and aggregate consequences. This paper explores these questions in the context of the rising spatial concentration of U.S. innovation, a striking trend in recent decades exemplified by the growth of high-tech clusters like Silicon Valley. The agglomeration benefits of these clusters have been widely discussed in urban economics and the popular press. Most notably, Moretti (2021) finds that individual inventors are more productive when located in regions with a high density of inventors, with agglomeration spillovers resulting in greater aggregate innovation than if inventors were evenly distributed across space. More fundamentally, however, little is known formally about why the growth of these high-tech clusters has been a relatively recent phenomenon in the United States. Do shifts in industry composition, firm entry, worker sorting by skill, or initial conditions explain the rise of high-tech clusters? Understanding the root causes driving the rising spatial concentration of innovation is crucial for developing effective place-based policies that might amplify these mechanisms and promote local growth. Additionally, the macroeconomic consequences of these high-tech clusters remain unclear. Do they boost the production of ideas that are adopted nationwide, thereby enhancing aggregate growth, or do these ideas remain confined within the high-tech clusters, potentially stifling broader economic development?

To answer these questions, it is essential first to carefully examine when and where the growth of high-tech clusters occurred. I begin by leveraging comprehensive data on patents and inventors to track how the geography of innovation in the United States has evolved over time. I find that the spatial concentration of patents remained relatively constant between 1976 and 1990 but increased dramatically from 1990 to 2018. Specifically, the Gini coefficient of patents per capita increased from approximately 0.36 in 1990 to 0.50 in 2018 – a change four times greater than the rise in income inequality in the US over the same period. Notably, this rising concentration primarily occurred in high-skill regions rather than in densely populated or large ones. The annual elasticity of patents per capita with respect to the Commuting Zone's (CZ) 1990 college ratio increased from about 1 in 1990 to 2 in 2018.

These facts suggest that a significant shock around 1990 triggered the growth of high-tech clusters. I show that the rapid rise of information and communication technologies (ICT) played two distinct roles in this process. First, the sudden arrival of influential ICT breakthroughs between 1985 and 1993 generated a sharp compositional shift in innovation toward ICT. The ICT share of all patents increased from roughly 8.5% in 1990 to 33% in 2018, and 78% of ICT patents during this period cited these breakthroughs within two degrees. This compositional shift matters for aggregate innovation concentration because ICT innovation is substantially more geographically concentrated than other fields: both the Gini coefficient of ICT patents per capita and the elasticity of CZ-level ICT patenting with respect to the 1990 college ratio are nearly twice as large as those for non-ICT patents. This heightened concentration reflects the longstanding *colocation* of ICT innovation and ICT production – encompassing Software Publishers; Telecommunications; and Data Processing, Hosting, and Related Services – which had already clustered in high-skill CZs before 1990 due to their deeper pools of high-skill labor. A decomposition of the aggregate trends in innovation concentration

by technology field shows that the rising ICT share accounts for about half of the total increase in innovation concentration in high-skill CZs between 1990 and 2018.

Second, ICT-induced reductions in communication costs beginning in the mid-1980s facilitated firm spatial expansion, which paradoxically concentrated innovation in high-skill CZs even further. Firms whose R&D was initially located in high-skill CZs (“high-skill CZ firms”) expanded into substantially more regions – both high- and low-skill – than firms with R&D in low-skill CZs and, following expansion, employed larger R&D workforces and generated more patents, the majority of which were in high-skill CZs. This phenomenon arises for both ICT and non-ICT oriented firms. I interpret it as an *asymmetric scale effect*: high-skill CZ firms can exploit lower production and service costs in other regions, whereas low-skill CZ firms already operate in relatively low-cost environments and therefore gain less from expansion. A firm-level decomposition of the post-1990 rise in ICT and non-ICT innovation concentration suggests that about 40% of the aggregate increase can be attributed to the growth of these high-skill CZ firms.

Finally, part of the growth of the ICT industry and of high-skill CZ firms – and hence aggregate innovation concentration – was driven by worker sorting and migration into high-skill CZs, which was most pronounced in the 1980s. Trends in the rolling-decade correlations of changes in the log college ratio and total worker population with the initial college ratio from 1970 to 2018 show that the correlation was roughly 0.10 in 1990, and about 0.02 in 2000 and thereafter.

To formalize how the ICT shock shaped the geography of innovation through the mechanisms above, I *introduce endogenous and directed innovation* into existing quantitative spatial models. Specifically, I develop a quantitative spatial growth model with endogenous innovation, technology diffusion, and dynamic worker mobility. Building on techniques from Eaton and Kortum (2001) and Lind and Ramondo (2024), I first integrate endogenous innovation and technology diffusion at the level of *individual ideas*, the fundamental unit in the Eaton-Kortum structure (Eaton and Kortum, 2024). This allows the model to characterize the degree of *colocation* between innovation and production and generate the *asymmetric scale effect*. I then introduce dynamic worker mobility, allowing workers to relocate in response to wage differences across regions and sectors and between production and research. Worker mobility amplifies both mechanisms and shapes the spatial direction of innovation following the ICT shock.

My model features two sectors, ICT and non-ICT. In each region and sector, innovation workers discover new ideas according to a Poisson process. Once an idea is discovered in a particular region, its diffusion to all regions is governed by independent Poisson processes with heterogeneous bilateral diffusion speeds. I capture the ICT shock through two exogenous components: (i) a one-time increase in the national productivity of ICT innovation, reflected as a rise in the Poisson arrival rate of new ideas in the ICT sector; and (ii) a decrease in the national component of bilateral diffusion speeds, representing reduced communication costs. Each idea corresponds to the production of a specific good on the unit interval and has two additional attributes: stochastic quality, drawn from a Pareto distribution upon discovery, and stochastic applicability, drawn from a separate Pareto distribution upon arrival in a new region. The productivity of an idea in producing its corresponding good is the product of its quality and applicability. This microfounded structure generates a

multivariate productivity distribution for goods at each instant. The marginal distribution in each region is Fréchet à la Eaton and Kortum (2002), but with an endogenous scale parameter determined by the history of innovation in all regions and diffusion speeds in all region-pairs. The equilibrium trilateral trade shares are determined by the product of idea adoption and idea market shares, capturing the two central mechanisms in my empirical analysis. Idea adoption shares govern the degree of *colocation* between innovation and production, while idea market shares generate the *asymmetric scale effect*: a uniform increase in bilateral diffusion speeds disproportionately raises the market access of ideas discovered in high-skill regions.

To enable these two mechanisms to shape the geography of innovation, I introduce *endogenous and directed innovation* through imperfect competition, which generates profits from production and thus incentives to innovate, together with worker mobility. In my model, each region-sector market contains a unit continuum of firms, each employing innovation workers and owning the ideas they create; each firm is therefore a collection of ideas. These firms engage in Bertrand competition à la Bernard et al. (2003): the lowest-cost producer of each good captures the entire market and charges the highest markup that deters any competitor from entering. This market structure generates an endogenous markup distribution across ideas, driven by their stochastic quality, while aggregate profits from sales in any destination market is a constant share of total income or expenditure in that market. The expected value of an idea in any destination market is the product of its market share and total profits earned there. Consequently, innovation wages are determined by the sum of an idea's expected values across all destination markets, multiplied by the Poisson arrival rate in the region where innovation occurs. Given wages from production and innovation, workers make dynamic decisions about moving across regions and sectors, as well as between production and innovation. Along the transition path, the ratio of real innovation wages across regions governs the incentives for worker mobility and therefore the spatial direction of innovation.

An additional advantage of my model is its ability to characterize how the geography of innovation affects aggregate growth and shapes the welfare impact of the ICT shock, thereby establishing a direct connection between quantitative trade and spatial models with innovation and endogenous growth models in macroeconomics. Along the balanced growth path, prices in all regions fall at the same aggregate rate, mirroring macroeconomic growth models. Regional wages differ but remain constant over time, determined by the spatial distribution of workers, trade, and technology diffusion like in quantitative trade and spatial models. Unlike macroeconomic growth models, however, the aggregate rate of price decline in my model is endogenously determined by the spatial distribution of innovation rates. And unlike trade and spatial models, a temporary shock to fundamentals affects not only the steady-state distribution of trade shares and nominal wages but also the long-run growth rate of the economy through falling prices. Using the characterizations of the balanced growth and transition paths, I then *analytically decompose the welfare impact* of the ICT shock – or any other shock to economic fundamentals – into its *transitory and long-run growth components*.

## Contributions to the Literature

This paper makes two main contributions across several literatures. First, I leverage comprehensive data on patents, inventors, and firms to examine when, where, and why innovation became increasingly spatially concentrated in the United States. While the empirical literature in innovation and urban economics has documented the rise of high-tech clusters in recent decades (e.g. Feldman and Kogler, 2010; Andrews and Whalley, 2022) and highlighted the benefits these clusters provide – such as enhancing inventor productivity and connecting innovation with academic science (e.g. Moretti, 2021; Bikard and Marx, 2020) – the more enduring and challenging question of what drives the growth of these high-tech clusters remains unanswered. I show that the post-1990 growth of high-tech clusters primarily occurred in high-skill CZs, and that the ICT shock likely drove this trend through three distinct mechanisms: *colocation*, the *asymmetric scale effect*, and *worker sorting*. The first mechanism – the colocation of ICT innovation and production – is a sector-specific extension of the broader colocation of innovation and production documented within the United States by Fort et al. (2020) and across countries by Liu (2024). The second mechanism highlights the geography of firm spatial expansion and its aggregate consequences on the spatial concentration of innovation. In particular, I show that firms initially concentrated in high-skill CZs expanded more geographically than other firms. This extends the empirical firm network literature, which documents the spatial expansion of firms (e.g. Hsieh and Rossi-Hansberg, 2023; Jiang, 2024; Kleinman, 2025) but largely abstracts from the geographic dimensions of this expansion. The third mechanism highlights the dynamics of worker sorting by education from 1970 to 2020, extending the cross-sectional comparisons between 1980 and 2000 or 2010 in Moretti (2013); Diamond (2016); Diamond and Gaubert (2022).

Second, I introduce *endogenous and directed innovation in a spatial setting*, thereby contributing to the trade, spatial, and macroeconomics literatures. To achieve this, I develop a model of spatial growth with two key ingredients. The first ingredient integrates endogenous innovation with technology diffusion at the level of *individual ideas*, advancing the quantitative trade and innovation literature. Specifically, unlike Eaton and Kortum (2001) and Lind and Ramondo (2023, 2024), innovation in my framework is fully endogenous and depends not only on equilibrium trade but also the entire technology diffusion network. Consequently, I characterize the degree of *colocation* between innovation and production and explain the *asymmetric scale effect*. These equilibrium results correspond to the spatial mechanics of innovation documented in my empirical findings and cannot be derived in existing trade and growth models, where innovation is either highly stylized or modeled as independent of technology diffusion. Most notably, Buera and Oberfield (2020) model technology diffusion as the transfer of knowledge from existing goods to the creation of new ones, such that the technology diffusion network does not impact profits from innovation. Somale (2021) incorporates endogenous innovation but excludes technology diffusion. Two recent papers integrate trade, innovation, and diffusion, albeit with simplifying assumptions. Cai et al. (2022) assumes perfect substitutability of ideas diffused from different locations, resulting in a scenario where small changes in relative wages cause large shifts in idea adoption and trade shares – a feature Eaton and Kortum term “the problem of flats”. Meanwhile, Xiang (2023) assumes instantaneous diffusion, limiting the model’s ability to capture the spatially heterogeneous effects of a uniform increase in bilateral diffusion speeds on idea market access.

The second ingredient of my model is dynamic worker mobility across regions and sectors and between production and research. This contrasts with the existing quantitative trade and innovation literature, which typically assumes perfect mobility restricted solely to movements between production and research within each region. With both ingredients, my model characterizes both the balanced growth path – the primary focus of this literature – and the transition path in response to shocks to economic fundamentals. Thus, my work aligns with the class of quantitative dynamic spatial models that incorporate worker migration (Caliendo et al., 2019), capital accumulation (Kleinman et al., 2023), and knowledge diffusion (Cai et al., 2025), all of which feature rich transition dynamics. I extend this class of models by introducing *endogenous and directed innovation* and by *decomposing the welfare effects of shocks to fundamentals into their transitory and long-run growth components*. Desmet et al. (2018) offers perhaps the only workhorse quantitative spatial model featuring endogenous innovation. However, they model endogenous innovation under perfect competition, where incentives to innovate arise from land rents. In contrast, my model fully microfounds innovation and technology diffusion *within* the Eaton-Kortum framework. Consequently, the spatial and sectoral directions of innovation on the transition path in my model depend not only on local population but also on equilibrium trade and technology diffusion networks. My analytical characterization of the spatial direction of innovation also advances the endogenous growth literature in macroeconomics, which has exclusively focused on the sectoral direction of technological change (e.g. Acemoglu, 1998, 2002, 2007) while abstracting from spatial considerations.

## 2 What Drives the Rising Spatial Concentration of US Innovation?

One of the most salient facts in the empirics of innovation is its rising spatial concentration in the United States, exemplified by the growth of Silicon Valley-like clusters. The fundamental drivers of this trend have, however, remained elusive despite more than two decades of research on the geography of innovation. Understanding why innovation became increasingly concentrated in space requires careful consideration of when and where it happened. In this section, I use the universe of patents, inventors, and patenting firms from 1976-2018 to (i) document the evolution of the spatial concentration of innovation and its underlying geography; and (ii) present decomposition exercises and descriptive micro-level evidence that suggest the key mechanisms behind the rising spatial concentration of innovation.

### 2.1 Data

I obtain the universe of patents produced between 1976 and 2022 from PatentsView, supplemented with bulk files from the US Patent and Trademark Office (USPTO). My sample includes all 3.70 million utility patents where at least one inventor lists a US address. Each patent contains extensive information, including inventor addresses, which typically reflect their home city and state. Using the Google Maps API, I geocode these addresses and map them to various spatial resolutions within the United States. I use these inventor locations to calculate annual patent counts for each region and derive measures of the spatial concentration of innovation across regions over time. Additionally, every patent is assigned a unique primary Cooperative Patent Classification (CPC) technology class. To analyze the role of compositional changes in innovative activity across fields, I map these classes to broader technology fields and subfields by adapting the field classification methodology developed by the World Intellectual Property Organization. Most patents also have one or more assignees, typically US firms, that hold ownership of the intellectual property. To examine the roles of compositional changes across firms and firm spatial expansion, I link these patent assignees to the universe of firms in the restricted US Census Longitudinal Business Database (LBD) using crosswalks provided by Kerr and Fu (2008) and Dreisigmeyer et al. (2018). Within firms, I assign inventors to regions in two alternative ways: the commuting zone (CZ) of their home city, and the nearest CZ where the firm has an establishment following Fort et al. (2020). I supplement the US Census LBD data with establishment-level data from Dun and Bradstreet's raw data files. I obtain the annual composition of workers by college education in each industry and county from IPUMS Current Population Survey (CPS) (Flood et al., 2025) and IPUMS National Historical Geographic Information System (NHGIS) (Schroeder et al., 2025). See Appendix A for more details.

### 2.2 Fact 1: Innovation became more concentrated in high-skill CZs after 1990

I begin by using the locational Gini index (Krugman, 1992) to measure the aggregate spatial concentration of innovation annually from 1976 to 2018. For each commuting zone (CZ), I calculate its share of total patents and share of total population in the US. I then rank CZs by their patent-to-population share ratio and plot the cumulative sum of patent shares against the cumulative sum of population shares to construct

the locational Gini curve. The locational Gini index, proportional to the area between this curve and the 45-degree line, measures the concentration of patents across CZs relative to population, with each CZ weighted by its population share. A value of zero indicates perfect equality while a value of one reflects perfect inequality<sup>1</sup>.

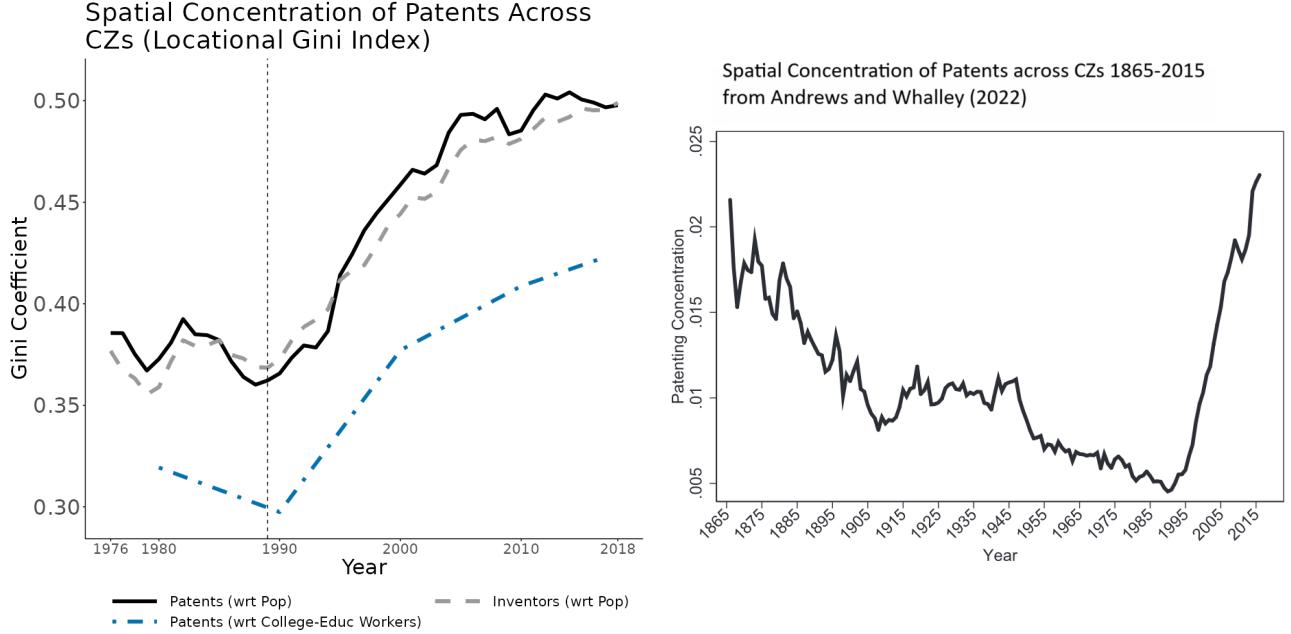


Figure 1: Trends in the spatial concentration of innovation across CZs. The left panel plots trends in the spatial concentration of innovation from 1976 to 2018, measured with respect to population and college educated workers using a locational Gini index (Krugman, 1991) and USPV data. For each CZ, I calculate annual shares of U.S. patents, inventors, population, and college-educated workers. Locational Gini curves are constructed for each year by ranking CZs based on the ratio of their patent or inventor share to their population or college-educated worker share, and plotting the cumulative sum of patent or inventor shares against the corresponding population or college-educated worker shares. The locational Gini index for each year is calculated as twice the area between the 45-degree line and the respective locational Gini curve. The right panel shows trends in the spatial concentration of innovation from 1865-2015 documented by Andrews and Whalley (2022) using a dartboard innovation intensity concentration index, a normalized measure of the sum of deviations of CZ patent shares from population shares.

Using this index, the solid black line in the left panel of Figure 1 shows that the spatial concentration of patents remained approximately constant from 1976 to 1990 but increased significantly from 1990 to 2018. The dotted grey line depicts the trend in the spatial concentration of inventors, computed annually by substituting patent shares with inventor shares in the locational Gini index. This trend closely mirrors that for patents, with the gap between the two from 1995 to 2018 potentially reflecting agglomeration economies in innovation, as documented by Moretti (2021)<sup>2</sup>. The dot-dashed blue line illustrates trends in the locational

<sup>1</sup>Formally, the locational Gini index  $G$  of patents  $x$  with respect to population  $L$  across all the  $N$  CZs is equivalent to half of the weighted mean absolute deviation of patents across all CZ-pairs  $o, d$ :

$$G = \frac{1}{2\mu} \sum_{o=1}^N \sum_{d=1}^N \frac{L_o}{L} \frac{L_d}{L} |x_o - x_d|, \quad \mu = \sum_{o=1}^N \frac{L_o}{L} x_o.$$

<sup>2</sup>Specifically, Moretti (2021) shows that inventors produce more patents when located in regions with more inventors. This

Gini index of patents relative to college-educated workers, computed annually by replacing population shares with college-educated worker shares in the locational Gini index. Relative to college-educated workers, the spatial concentration of patents remained approximately constant in the 1980s and rose significantly after 1990. This finding suggests that the rising spatial concentration of innovation cannot be solely attributed to the geographical sorting of workers by skill from 1980-2000. In Figure 11 in Appendix B.1, I demonstrate the robustness of these trends to: (i) excluding top patenting CZs such as San Jose, San Francisco, Newark, and Los Angeles, and; (ii) employing alternative measures of spatial concentration, including the coefficient of variation, Herfindahl index, simplified Ellison and Glaeser (1997) index, and annual share of patents produced by the top 10 CZs. The right panel presents trends in the spatial concentration of innovation from 1865 to 2015, as documented by Andrews and Whalley (2022) using an alternative patent dataset and a dartboard innovation intensity concentration index: a normalized measure of the sum of deviations of CZ patent shares from population shares building on the Ellison-Glaeser index. Consistent with my measure, this figure reveals a sharp rise in the spatial concentration of innovation beginning in 1990.

The locational Gini curve also provides a natural, annual measure of patenting activity for each CZ that allows for meaningful comparisons over time: the patent-to-population share ratio. This measure reflects patents per capita while normalizing the total number of patents and population in each year, effectively capturing each CZ's contribution to the overall locational Gini coefficient in that year. Compared to existing measures in the literature, changes in the patent-to-population share ratio over time reveal whether patents are increasingly concentrated in a given CZ, independent of scale and after controlling for changes in local and national populations<sup>3</sup>. In Appendix B.1, Figure 12 depicts the geography of changes in the patent-to-population share ratio between 5-year averages around 1990 and 2015 while Table 4 lists the top 15 CZs with the largest increase in this period. Apart from well-known superstar cities such as San Jose, San Francisco, San Diego, Seattle, and Boston, regions like Portland, Boise, Wayne, Provo, and Fort Collins also emerged as some of the most innovative areas in the United States between 1990 and 2015.

Using this measure, Figure 2 shows that innovation became increasingly concentrated in high-skill CZs between 1990 and 2018. The left graph plots the correlation between the log college ratio in 1990 and the percentage change in the patent-to-population share ratio from 1990 to 2015 across CZs. Specifically, a 1% increase in the college ratio in 1990 is associated with a 0.8% greater increase in the patent-to-population share ratio over this period. To account for regions with zero patents and to exploit the annual frequency of the dataset, I estimate annual elasticities  $\alpha_t$  of patents per capita<sup>4</sup> with respect to the 1990 college ratio – henceforth “patent elasticity” – using the following Poisson Pseudo Maximum Likelihood (PPML)

---

finding implies that the spatial concentration of patents exceeds that of inventors.

<sup>3</sup>Raw patent counts, commonly used in other papers, suffer from several limitations. First, they inherently favor regions with larger populations, leading to biased rankings that can change arbitrarily when regions are grouped differently. Second, intertemporal comparisons are less informative because increases in raw patent counts can result from aggregate growth in patenting, changes in a CZ's share of annual patents, or regional population growth. R&D expenditures, an alternative measure of innovation, are typically available only at the state level.

<sup>4</sup>This is equivalent to the patent-to-population share ratio with year fixed effects.

specification:

$$\text{Patents per capita}_{r,t} = \exp(\alpha_t \cdot \text{Log 1990 College Ratio}_r \times \text{Year}_t + \gamma_t + \epsilon_{r,t}). \quad (1)$$

where  $r$  represents regions and  $t$  denotes years. The estimated annual aggregate elasticities  $\alpha_t$  serve an important role for the remainder of my empirical analysis. The right graph of Figure 2 shows that this elasticity remained fairly constant at around 1 from 1976 to 1990 but rose sharply thereafter. This trend reflects the gradual geographical sorting of patents per capita across CZs, driven by differences in their initial college ratios in 1990. In Appendix B.1, Figure 14 plots trends in the elasticity of CZ patents per capita with respect to the 1990 population and population density, respectively. The absence of a clear break in these trends after 1990 suggests that the skill mix of workers in 1990 is a more important margin for the geographical sorting of patents per capita over time than either population size or density.

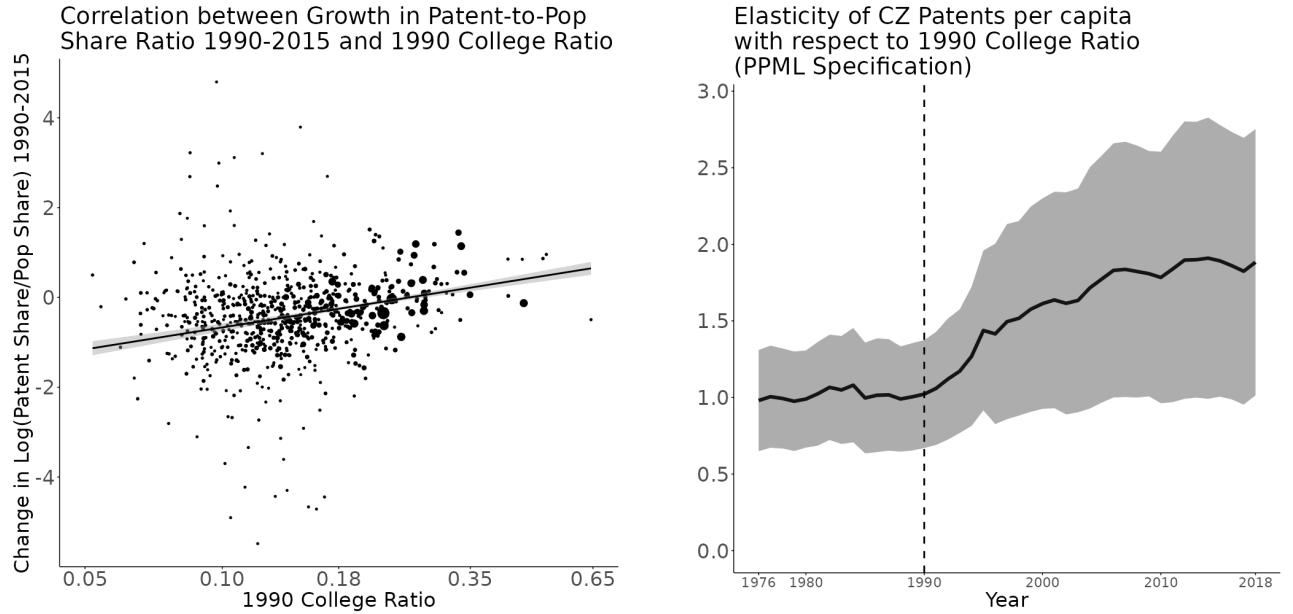


Figure 2: The post-1990 geographical sorting of patents per capita toward high-skill CZs. The left graph shows how percentage changes in the patent-to-population share ratio from 1990 to 2018 relate to the initial college ratio across CZs. Specifically, it plots changes in log patent-to-population share ratio between five-year averages around 1990 and 2015 against the 1990 college ratio on a log scale, with the size of each CZ's dot representing its population. The right graph displays trends in the annual elasticity  $\alpha_t$  of CZ patents per capita with respect to the 1990 college ratio (henceforth the “patent elasticity”). The confidence intervals in both graphs reflect heteroskedastic-robust standard errors.

## 2.3 An Anatomy of the Post-1990 Rise in the Spatial Concentration of Innovation through Proximate Field and Firm Decompositions

There are many potential explanations for the post-1990 rise in innovation concentration in high-skill CZs (**Fact 1**). To isolate the main mechanisms, I first decompose the aggregate trends by technology field and by firm. The results show that the surge in concentration was jointly driven by patent growth in *ICT-oriented*

firms and in *high-skill CZ firms*, i.e. ICT and non-ICT firms that concentrated their *innovation* in high-skill CZs pre-1990. Section 2.4 then provides descriptive micro-level evidence on the mechanisms underlying patent growth in these two groups of firms.

### 2.3.1 Field Decomposition

I begin with the field decomposition. The left panel of Figure 3 displays the evolution of each field's share of annual patents, while the right panel depicts the corresponding field-specific patent elasticity ( $\alpha$ ). Along with the Shapley decompositions reported in Appendix B.2, these patterns suggest two main channels behind the post-1990 surge in patent concentration in high-skill CZs, discussed in the next two facts.

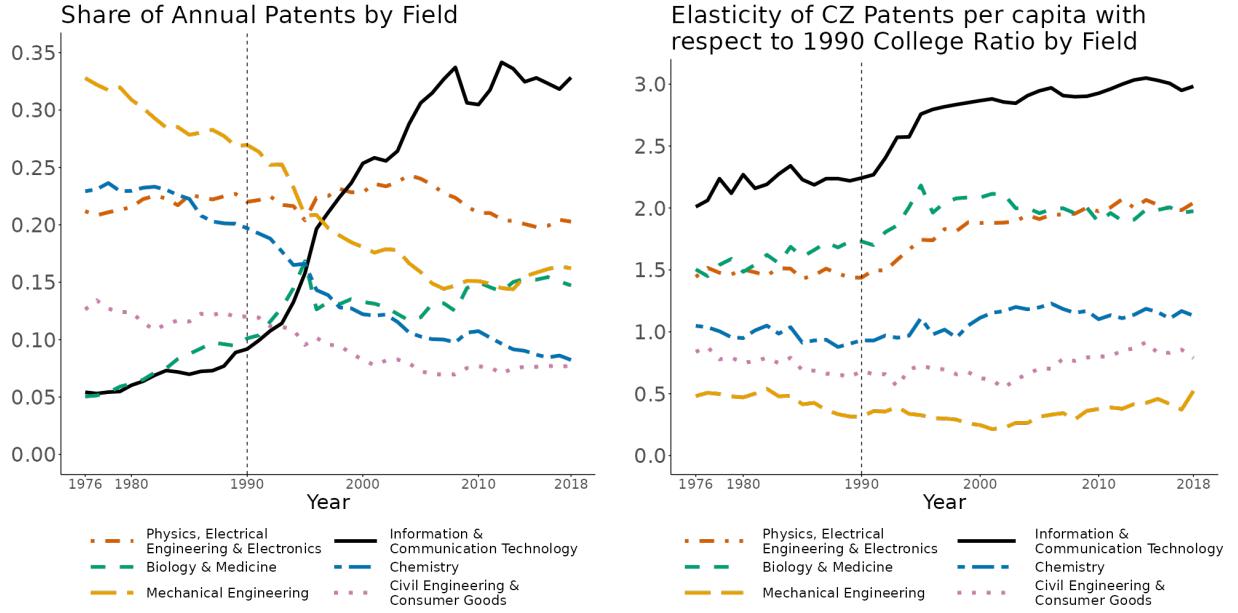


Figure 3: Trends by technology field: annual patent shares (left panel) and the elasticity of patents per capita with respect to the 1990 college ratio, estimated using PPML (right panel)

#### **Fact 2: About half of the post-1990 increase in concentration reflects a compositional shift in innovation toward ICT, a field more concentrated in high-skill CZs**

The share of ICT patents has risen steadily since the late 1980s. The left panel of Figure 3 shows that ICT patents increased from about 8% in 1990 to over 32% by 2015. Because ICT innovation is systematically more concentrated in high-skill CZs than non-ICT innovation, this reallocation of innovative activity toward ICT substantially raised aggregate concentration. The right panel of Figure 3 shows that the patent elasticity for ICT is consistently higher than that of other fields. A formal Shapley decomposition (Appendix B.2.2) attributes 53% of the post-1990 increase in aggregate concentration to this between-field reallocation toward ICT<sup>5</sup>. At the subfield level (Appendix B.2.6), the compositional shift is driven almost entirely by the growing shares of *Digital Communications* and *Computer Technology*. These trends and results are robust to alternative specifications, including dropping top patenting CZs (Appendix B.2.5) and measuring

<sup>5</sup>Note that an additional 7% reflects the rising share of Biology patents

spatial concentration using the Gini coefficient of patents per capita rather than restricting to high-skill CZs (Appendix B.2.3).

**Fact 3: Beyond composition, both ICT and non-ICT innovation became increasingly concentrated in high-skill CZs after 1990**

The right panel of Figure 3 also shows that patents in most fields became more concentrated in high-skill CZs after 1990, as reflected in rising field-specific elasticities. The field decomposition in Appendix B.2.2 indicates that 40% of the aggregate increase in concentration reflects rising elasticities within fields, with roughly two-thirds of this contribution coming from non-ICT fields. A finer decomposition by subfield (Appendix B.2.6) shows that the within-field rise is broadly distributed across many subfields, in sharp contrast to the compositional shift, which was driven almost entirely by *Digital Communications* and *Computer Technology*. These results show that the aggregate increase in concentration cannot be attributed solely to patenting in the ICT sector.

### 2.3.2 Firm Decomposition

To better understand the forces behind the compositional shift toward ICT and its greater concentration in high-skill CZs (Fact 2), as well as the rise in within-field concentration across ICT and non-ICT (Fact 3), I further decompose each channel into within-firm and between-firm components.

**Fact 4: The post-1990 shift in innovation toward ICT was driven by the growth of *ICT-oriented firms*, which disproportionately concentrate their patents in high-skill CZs**

The first channel – the compositional shift toward ICT and its greater concentration in high-skill CZs – is primarily driven by the rising patent shares of *ICT-oriented firms*.

ICT-oriented firms drive the compositional shift toward ICT. The left panel of Figure 4 presents results from a modified Foster et al. (2001) [henceforth FHK] decomposition<sup>6</sup> of the post-1990 rise in the aggregate patent share (see Appendix B.3.1 for details). This increase (solid black line) is explained almost entirely by the between-assignee component (dot-dashed blue line): the ICT share within individual assignees remains relatively stable over time, while ICT-oriented firms expand their overall share of patenting.

ICT-oriented firms disproportionately patent in high-skill CZs, for both ICT and non-ICT technologies<sup>7</sup>. The middle panel of Figure 4 shows weighted averages of firm-level elasticities for ICT-oriented and non-ICT firms (see Appendix B.3.1 for the precise definitions). Non-ICT patents produced by ICT-oriented firms (dashed orange line) are more concentrated in high-skill CZs relative to ICT patents produced by non-ICT-oriented firms (dotdashed blue line). The right panel of Figure 4 formalizes this comparison using a Shapley

---

<sup>6</sup>I use the FHK decomposition, instead of a Shapley decomposition, to appropriately account for firm entry and exit.

<sup>7</sup>Although the compositional shift toward ICT is largely explained by the rising patent shares of ICT-oriented firms, aggregate differences between ICT and non-ICT elasticities could, in principle, reflect either cross-firm differences in patent geography or within-firm differences between ICT and non-ICT patents. For example, both ICT- and non-ICT-oriented firms might place their ICT patents disproportionately in high-skill CZs relative to their non-ICT patents.

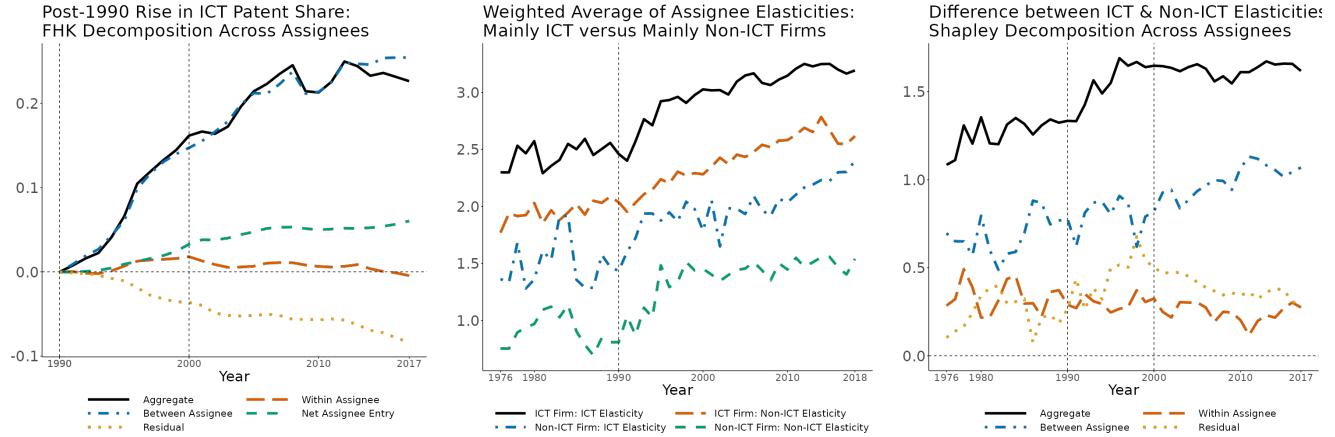


Figure 4: Decomposition of the aggregate rise in the ICT patent share into within versus across firm (left) and into firms that started patenting before versus after 1990 (right).

decomposition<sup>8</sup> of the aggregate difference between ICT and non-ICT elasticities by year (see Appendix B.3.1 for details and robustness checks). The aggregate difference (solid black line) consistently arises more from the between-assignee component (dotdashed blue line) than from the within-assignee component (dashed orange line), reflecting systematic differences in the geography of patenting between ICT and non-ICT-oriented firms, rather than differences in the geography of ICT and non-ICT patents within firms.

#### Fact 5: The post-1990 increase in the spatial concentration of ICT and non-ICT innovation was driven by the growth of *high-skill CZ firms*

The second channel – the rise in patent elasticities within fields – is primarily driven by the growing patent share of *high-skill CZ firms*, i.e. firms already concentrated in high-skill CZs.

Figure 5 presents results from a modified FHK decomposition of the post-1990 increase in the aggregate spatial concentration of ICT and non-ICT patents in high-skill CZs (see Appendix B.3.3 for details). The left panel shows that two-thirds of the post-1990 rise in the ICT patent elasticity (solid black line) is explained by the between-assignee component (dotdashed blue line), reflecting the rising patent share of firms that already exhibited higher ICT elasticities. The right panel shows that virtually all of the post-1990 increase in the non-ICT patent elasticity is accounted for the between-assignee component, reflecting the rising patent share of firms that already exhibited higher non-ICT elasticities. Appendix B.3.3 shows that these patterns reflect the rising patent shares of both ICT- and non-ICT-oriented firms with high initial elasticities, rather than a compositional shift toward ICT-oriented firms or the entry of new firms.

#### 2.4 Descriptive Evidence on the Underlying Mechanisms

The field and firm decompositions (**Facts 2 to 5**) illustrate that the post-1990 rise in the spatial concentration of innovation was driven by the growth of two groups of firms: *ICT-oriented* and *high-skill CZ*. This

<sup>8</sup>Since the decomposition is conducted separately for each year, firm entry and exit do not pose an issue, so an FHK decomposition is unnecessary.

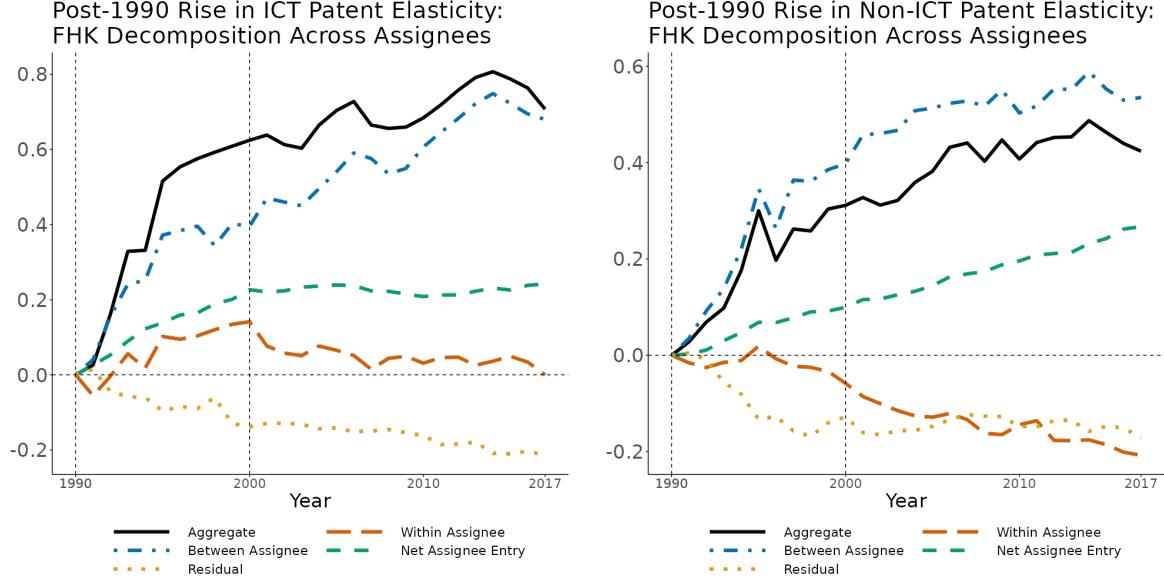


Figure 5: Trends in the decomposition of the post-1990 rise in the aggregate ICT (left panel) and non-ICT (right panel) innovation concentration in high-skill CZs.

raises two questions: What drove the post-1990 expansion of these firms? And why do ICT-oriented firms locate a larger share of their patents in high-skill CZs relative to non-ICT-oriented firms?

I show below that their growth was likely shaped by two components of the ICT shock: (i) the sudden emergence of influential ICT breakthroughs, and (ii) the subsequent decline in communication costs enabled by ICT adoption. To illustrate the mechanisms, consider *ICT-oriented firms* such as IBM and Hewlett-Packard (HP) alongside *non-ICT high-skill CZ firms* such as Boston Scientific, United Technologies, Boeing, and Exxon Mobil. Before 1985, IBM concentrated both production and R&D in the highly educated Northeast corridor, with additional facilities in Boulder, CO; Rochester, MN; and Silicon Valley. HP located nearly all of its manufacturing, engineering, and managerial capacity in Silicon Valley. In both cases, innovation was colocated with production in high-skill CZs to exploit deep local pools of engineers and technical workers essential for ICT development. The sudden arrival of influential ICT breakthroughs in the mid-1980s triggered sharp increases in patenting at IBM and HP, much of which remained colocated near their existing production sites. Beginning in the late 1980s – and facilitated by falling communication costs – IBM and HP, as well as high-skill CZ firms more broadly, began expanding service, engineering, and support operations across the country while simultaneously increasing their R&D employment and patenting activity. United Technologies, for example, broadened its footprint beyond its Connecticut core into engineering hubs such as Phoenix and Dallas–Fort Worth, as well as into lower-skill regions across the Southeast and Midwest where it added manufacturing and service operations. Exxon Mobil expanded from its traditional bases in Houston and Northern Virginia into both high-skill technical regions and lower-skill Gulf Coast and Mid-South locations, increasing its presence in refining, petrochemicals, and business services. Boston Scientific and Boeing exhibit similar patterns of mixed high-skill and low-skill expansion as their national production and service networks scaled. Across all of these cases, firms experienced pronounced surges in R&D and

patenting as they expanded geographically.

#### 2.4.1 ICT Innovation Breakthroughs and the *Colocation* of ICT Innovation with Production in High-Skill CZs

The first component of the ICT shock is the sudden emergence of influential ICT breakthroughs. I identify breakthrough ICT patents as those in the top 1% of the *originality* distribution across all patents, where originality is measured as the ratio of forward to backward textual similarity following Kelly et al. (2021). Between 1976 and 2018, 880,927 ICT patents were produced, of which 27,419 qualify as breakthroughs. The left panel of Figure 6 traces the evolution of their influence – defined as the number of patents that cite each breakthrough as their first breakthrough, following citations up to two degrees. The figure reveals a sharp rise in the influence of breakthrough ICT patents between 1984 and 1993. These breakthroughs catalyzed the subsequent surge in ICT patenting. The right panel of Figure 6 displays the corresponding trend in the number of ICT patents that cite these breakthroughs (within two degrees). Across all years, 78% of all ICT patents cite a breakthrough ICT patent within two degrees.

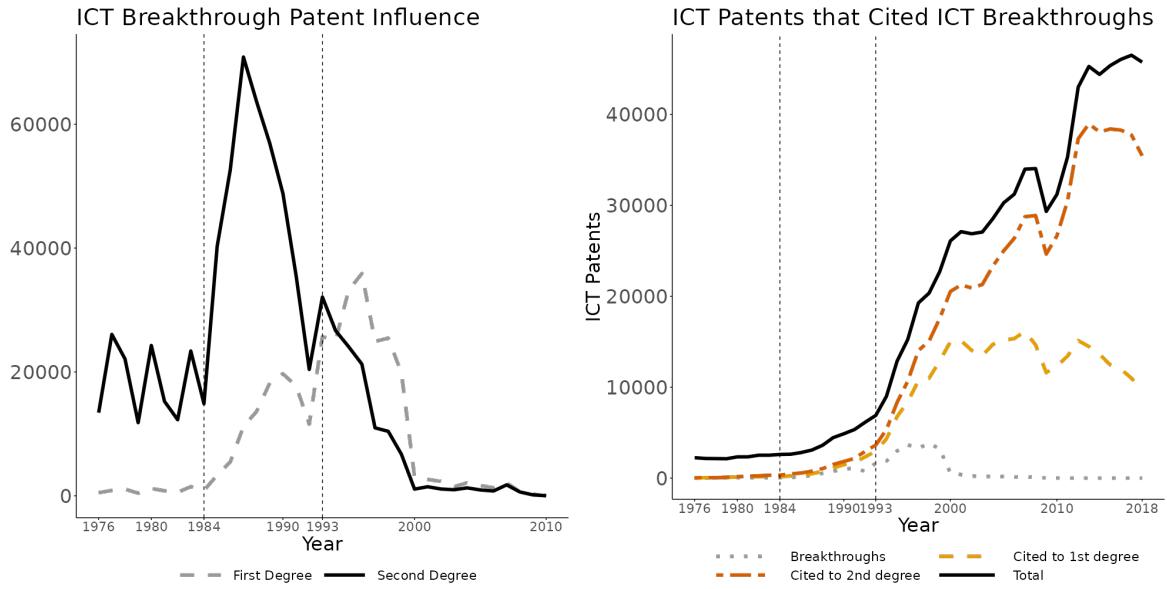


Figure 6: Origins of the ICT shock. The left panel shows the influence of these breakthrough patents, defined as the number of patents that cite each breakthrough as their first breakthrough, tracing citations up to two degrees. The right panel depicts trends in the number of ICT patents that cited these patents.

This surge in ICT innovation productivity disproportionately increased innovation in *ICT-oriented firms*, rather than uniformly raising ICT patenting across all firms, as shown in **Fact 4**. Most ICT-oriented firms operate plants in the following industries in the Information Sector (NAICS 51): Software Publishers (5112); Telecommunications (517), and; Data Processing, Hosting, and Related Services (518)<sup>9</sup>. This suggests that having an ICT service plant is a key input in the production of ICT innovation.

<sup>9</sup>In contrast to Fort et al. (2020), I categorize these industries as *ICT production*, rather than *innovation*, because establishments that focus on R&D activities are instead classified as Scientific Research and Development Services (5417) and occasionally Corporate, Subsidiary, and Regional Managing Offices (551114) under NAICS.

ICT service plants disproportionately employ high-skill workers and are therefore heavily concentrated in high-skill CZs. The left panel of Figure 7 shows that the share of college-educated workers in the ICT service industry (solid black line) is substantially higher than that in other industries (dashed blue line) across all years. The right panel of Figure 7 plots the relationship between a CZ's 1990 college ratio and its pre-1985 ICT employment share; the positive slope indicates that high-skill CZs already had disproportionately large ICT service employment before 1985.

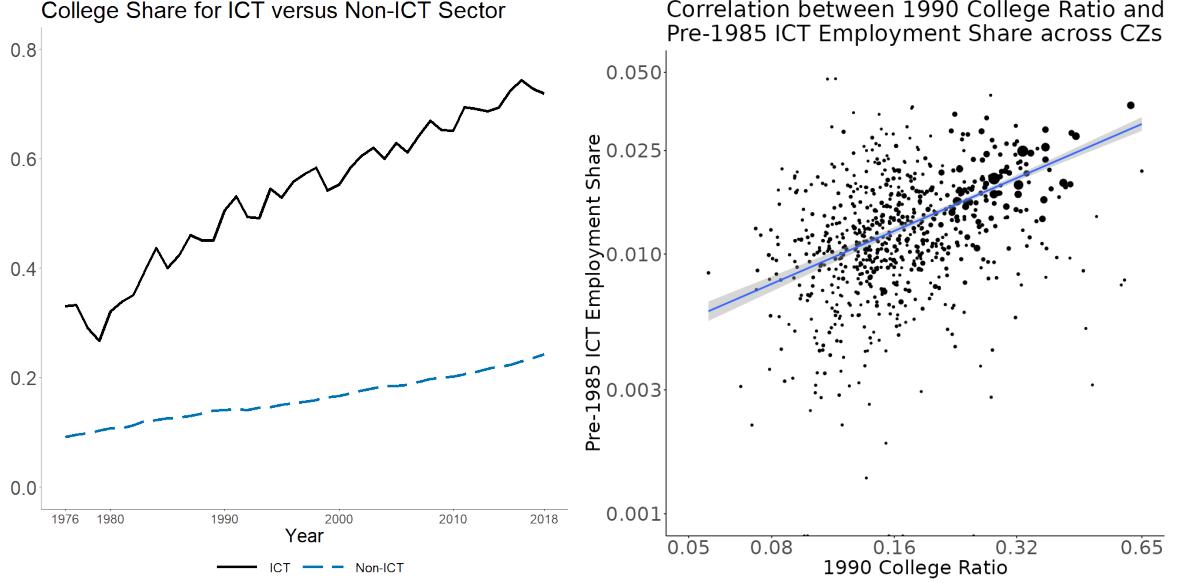


Figure 7: The ICT sector disproportionately employs college-educated workers and is concentrated in high-skill CZs. The left panel displays trends in the college share for the ICT and non-ICT sectors. The right panel shows the correlation of the pre-1985 ICT employment share against the 1990 college ratio. Each dot represents a commuting zone (CZ), where the size of the dot captures the CZ's population in 1990.

The growth of ICT patents primarily occurred in CZs where the ICT industry was initially more concentrated. To show this, I estimate the following CZ-year specification, reminiscent of the first stage of a Bartik regression:

$$\text{Outcome}_{r,t} = \beta \cdot \text{Pre-1985 ICT Employment Share}_r \times \text{Growth of National ICT Patents}_t + \gamma_r + \gamma_t + \varepsilon_{r,t} \quad (2)$$

where  $\text{Growth of National ICT Patents}_t$  is the growth of national patents in year  $t$  relative to 1985, and  $\gamma_r$  and  $\gamma_t$  are CZ and year fixed effects. Specification (1) shows that CZs with a higher initial ICT employment share experienced substantially larger increases in ICT patenting. A 1 p.p. higher pre-1985 ICT employment share is associated with 8.3 more ICT patents in years when national ICT patenting is 10% above its 1985 level. Specification (2) confirms that the geographical incidence of ICT patents is not driven merely by population growth: the same 1 p.p. difference predicts 1.5 additional ICT patents per thousand residents under a 10% national ICT growth shock. Turning to employment outcomes, Specification (3) shows that ICT employment also increased disproportionately in CZs where the ICT industry was initially more concentrated: a 1 p.p. higher pre-1985 ICT employment share is associated with 77 additional ICT workers when national ICT patenting rises by 10%. Taken together, the parallel rise in ICT employment and ICT patents per

capita in the same CZs provides evidence of the colocation of ICT innovation and ICT production in regions with strong pre-existing ICT presence. Finally, Specification (4) shows that these CZs did not experience larger increases in the share of local employment in ICT, likely because employment in other industries also expanded. I return to this result when discussing worker migration in Section 2.4.3.

Dependent Variables:	ICT Patents	ICT Patents per capita	ICT Employment	ICT Employment Share
Model:	(1)	(2)	(3)	(4)
Pre-1985 ICT Employment Share <sub>r</sub> × Growth of National ICT Patents <sub>t</sub>	8,334.9*** (2,735.3)	1.474*** (0.5416)	77,372.4*** (28,464.3)	-0.0247 (0.0204)
<i>Fixed-effects</i>				
CZ	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Observations	20,938	20,938	20,938	20,934
R <sup>2</sup>	0.81282	0.81900	0.96195	0.83343
Within R <sup>2</sup>	0.15934	0.06459	0.05407	0.01308

*Clustered (CZ) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 1: The geographical incidence of the post-1985 growth of national ICT patents.

#### 2.4.2 Reduced Communication Costs and *Firm Spatial Expansion*

Table 2 lists the top five breakthrough ICT patents between 1984 and 1993 by second-degree citation influence. Although most of these breakthroughs reflect advances in the application or administrative layers of ICT, patent US4644532 stands out for improving distributed routing in networks, indirectly lowering communication costs and enabling more robust distributed information systems.

Beyond the sudden arrival of these breakthrough patents, the ICT revolution was also propelled by the *adoption* and *standardization* of foundational digital technologies. The transition from proprietary systems to open standards – Ethernet, TCP/IP, relational databases, and client–server architectures – occurred primarily during the early-to-mid 1990s, when firms finally achieved seamless interconnection across plants, offices, and data systems (Bresnahan and Greenstein, 1996; Mowery and Simcoe, 2002). In parallel, the deployment and rapid scaling of the National Science Foundation Network (NSFNET) between 1986 and 1995 dramatically increased the availability of high-speed Internet access, further reducing coordination costs (Greenstein and Prince, 2006; Greenstein, 2015). See Appendix B.4 for more details.

As communication costs fell with the ICT revolution, firms expanded across many regions within the United States. This pattern is well documented in prior work: Jiang (2024) shows that the geographic scope of multi-unit manufacturing firms widened gradually from 1982 to 1997 and more rapidly from 1997 to 2007, with firm-level intranet adoption by 1999 causally linked to spatial expansion between 1997 and 2002. Kleinman (2025) finds a steady rise in the number of establishments operated by service firms from 1980 to 2017, accounting for most of these firms’ post-1990 growth. Hsieh and Rossi-Hansberg (2023) document similar nationwide patterns: a modest increase in the number of MSAs per firm from 1977 to 1987, followed by a

Patent No.	Year	Assignee	Title	2nd Deg.	1st Deg.
4827508	1985	Personal Library	Database usage metering and protection system and device	5135	671
5172338	1990	Sandisk	Multi-state EEPROM read and write circuits and methods	3771	639
<b>4644532</b>	<b>1985</b>	<b>IBM</b>	<b>Automatic update of topology in a hybrid network</b>	<b>3571</b>	<b>223</b>
5005122	1987	DEC/HP	Arrangement with cooperating management server node and network service node	2969	1068
4831526	1986	Chubb	Computerized insurance premium quote request and response system	2528	333

Table 2: Top 5 Breakthrough ICT Patents from 1984-1993 by Second-Degree Citation Influence

sharper rise thereafter – likely reflecting the diffusion of ICT technologies and complementary management practices. Using the raw D&B files, I find closely related trends in my data. The left panel of Figure 8 shows that the employment and sales shares of multi-CZ firms increased from about 94% to 96.5% between 1990 and 2018. The right panel shows that the employment- and sales-weighted mean number of CZs per firm rose from roughly 100 to 180 over the same period.

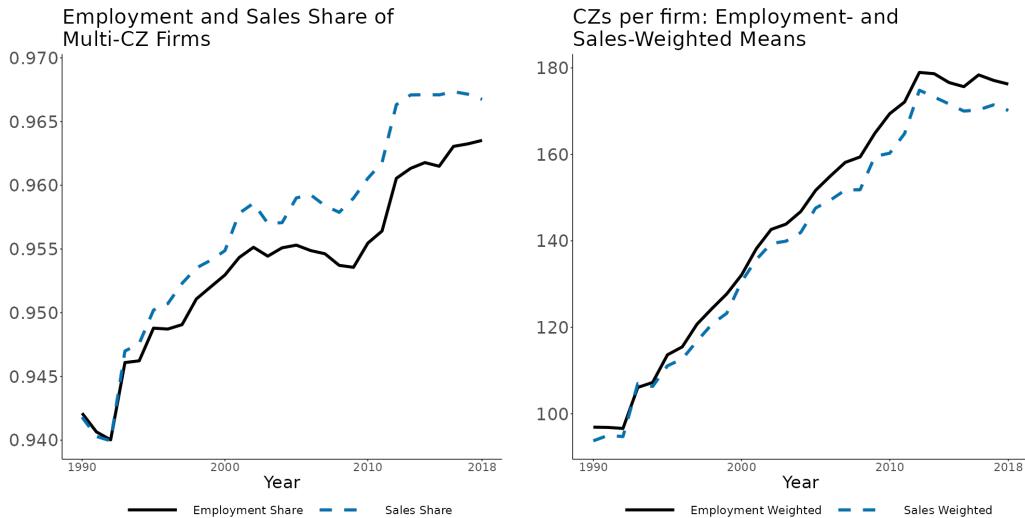


Figure 8: Trends in the *intensity* of firm spatial expansion from 1990 to 2018. The left panel displays trends in the employment and sales share of firms with establishments in multiple CZs. The right panel shows trends in the employment- and sales-weighted mean number of CZs per firm.

I relate firm spatial expansion to the post-1990 rise in the concentration of innovation toward high-skill CZs. In Figure 9, the left panel shows that firms with initial R&D capabilities – defined as those operating a plant in Professional, Scientific, and Technical Services (NAICS 54) or Management of Companies and Enterprises (NAICS 55) between 1990 and 1992 – had a broader geographic footprint in 1990 and expanded more uniformly across high-skill and low-skill CZs thereafter than firms without such capabilities. The right panel shows that firms with R&D operations in high-skill CZs – where over 90% of all R&D-plant employment is located in every year – expanded geographically far more than firms with R&D plants in

low-skill CZs between 1990 and 2005.

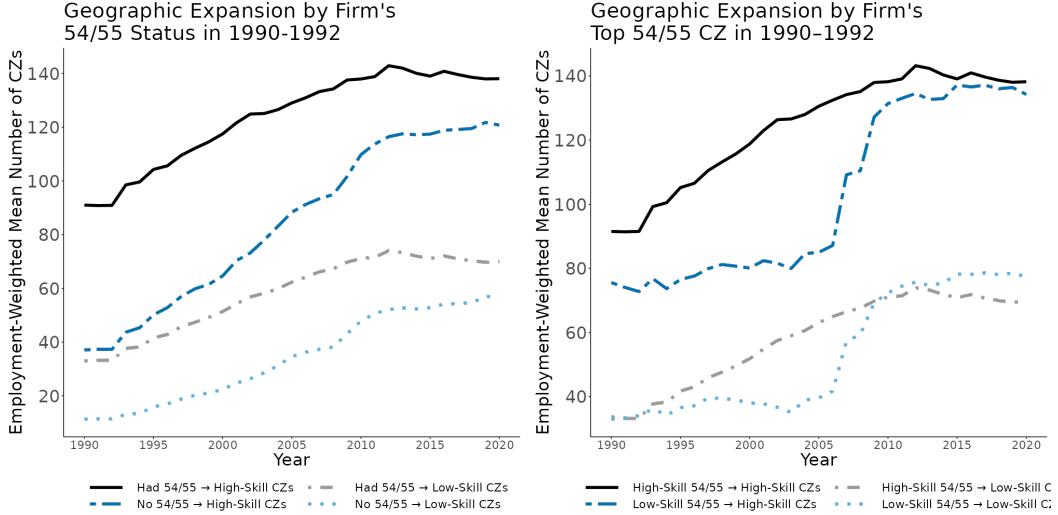


Figure 9: Trends in the *geography* of firm spatial expansion from 1990 to 2018. The left panel displays employment-weighted trends in the mean number of high-skill and low-skill CZs served by firms with an R&D plant in 1990–1992 versus those without one. The right panel shows analogous trends for firms with R&D plants in high-skill CZs versus firms with R&D plants in low-skill CZs.

Importantly, firms that expanded into more CZs subsequently increased their R&D activity in high-skill CZs. Table 3 indicates that expanding a firm’s geographic footprint by 10 CZs is associated with a 0.6-percentage-point increase in the likelihood of having an R&D plant (Specification 1) and a 1.1-percent increase in R&D employment (Specification 2) in the following year. Specification (3) shows that virtually all of this additional R&D employment occurred in CZs in the top skill quartile. This pattern complements the findings of Hsieh and Rossi-Hansberg (2023), who document that much of the growth in markets per firm occurred in finance, retail, and other services, where leading firms expanded into smaller MSAs and simultaneously increased employment in their R&D establishments.

Dependent Variables:	Has R&D Extensive Margin (1)	Log R&D Employment Intensive Margin (2)	Log R&D Employment (Q4 CZs) Intensive Margin (3)
Model:			
<i>Fixed-effects</i>			
Firm	Yes	Yes	Yes
Year	Yes	Yes	Yes
<i>Fit statistics</i>			
Observations	993,573	28,605	28,544
R <sup>2</sup>	0.81316	0.83793	0.85087
Within R <sup>2</sup>	0.00386	0.02642	0.02564

*Clustered (firm) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 3: Correlation of Lagged Firm Geographic Expansion on R&D Activity

### 2.4.3 Worker Sorting to High-Skill CZs

Part of the growth of ICT-oriented firms and firms in high-skill CZs was supported by *worker sorting* toward these locations. Specifications (3) and (4) in Table (2) show that employment in both ICT and non-ICT sectors rose disproportionately in high-skill CZs after 1985, indicating that this expansion was fueled by in-migration of workers. A well-established finding in spatial economics is that U.S. workers increasingly sorted by education level after 1980 (Moretti, 2013; Diamond, 2016; Diamond and Gaubert, 2022). Whereas prior studies have focused on cross-sectional comparisons between 1980 and 2000 or 2010, I examine the dynamics of this process over time to understand how worker sorting relates to the ICT shock and the subsequent rise in the spatial concentration of innovation. Figure 10 plots rolling-decade correlations of changes in the log college ratio (left panel) and changes in total worker population (right panel) with the initial college ratio from 1970 to 2018. The results show that most of the sorting and migration toward high-skill CZs occurred during the 1980s – before the post-1990 increase in innovation concentration – though the process continued, more gradually, in subsequent decades. This pattern suggests that the inflow of high-skill workers during the 1980s helped create the human-capital foundations that enabled the subsequent growth of ICT-oriented and high-skill-CZ firms after 1990.

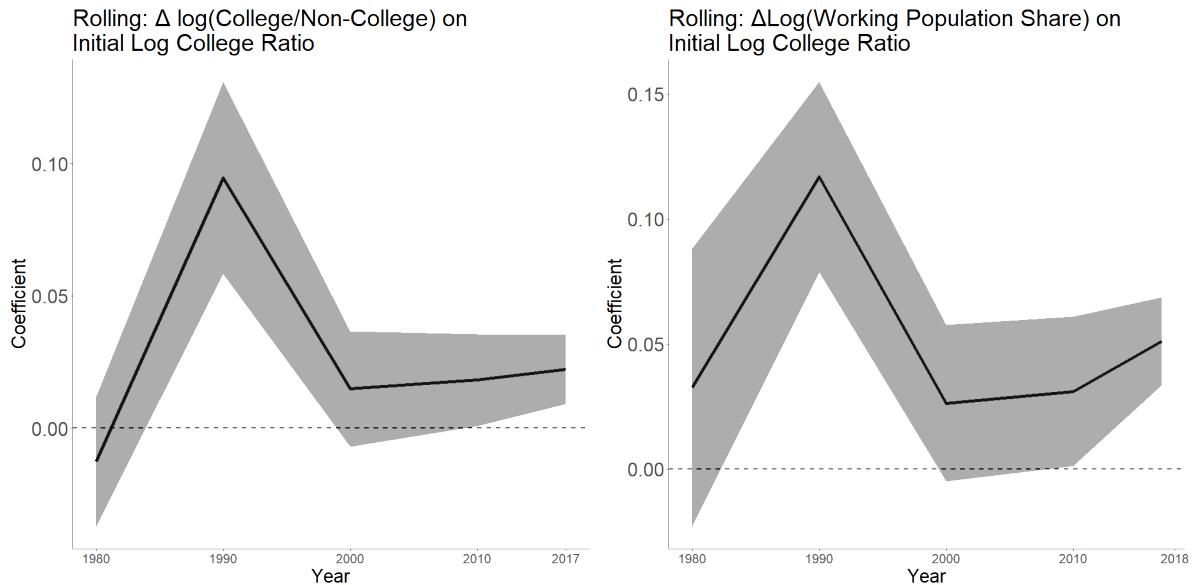


Figure 10: Worker sorting and migration to high-skill CZs from 1980 to 2018. Trends in the rolling correlation of changes in the log college ratio (left panel) and in worker population (right panel) with respect to the initial college ratio for each decade from 1970 to 2018.

### 3 A Model of Spatial Growth with Endogenous and Directed Innovation, Technology Diffusion, and Dynamic Worker Mobility

My empirical findings suggest that the post-1990 rise in the spatial concentration of innovation in high-skill CZs (Fact 1) is primarily driven by three mechanisms: (i) the growth of the ICT sector – whose production and innovation are *colocated* in high-skill CZs – following the emergence of influential ICT breakthroughs; (ii) *firm spatial expansion* enabled by falling communication costs; and (iii) *worker sorting* toward high-skill CZs.

To decompose the role of each mechanism and assess the aggregate consequences of the rising spatial concentration of innovation, I develop a quantitative spatial growth model with two main components. First, I integrate endogenous innovation and technology diffusion at the level of *individual ideas*, the fundamental unit of the Eaton-Kortum world. Leveraging tools from Eaton and Kortum (2024), I introduce imperfect competition and endogenous innovation in Lind and Ramondo (2024)'s quantitative trade model with idea diffusion. This structure endogenously generates the colocation of ICT innovation with production following a positive shock to ICT research productivity and captures how firm spatial expansion, enabled by falling communication costs, influences profits through technology diffusion and thereby shapes the geography of innovation. Second, I incorporate dynamic worker mobility across regions, sectors, and between production and research, extending Caliendo et al. (2019). This component captures how worker sorting facilitates the above mechanisms and, in turn, drives the spatial distribution of innovation.

More broadly, the model delivers two methodological contributions. First, it *introduces directed innovation<sup>10</sup> in a quantitative spatial model*, allowing spatial patterns of innovation to evolve endogenously and gradually in response to the ICT shock through the mechanisms identified in the empirics. Second, it provides an *analytical decomposition of the welfare effects of any shock to economic fundamentals into a transitory component and a long-run growth component*. The decomposition, developed in the next section, illustrates how the ICT shock – through its impact on the geography of innovation – shapes aggregate growth and welfare.

Formally, the economy consists of  $N$  regions, indexed by  $r,o,d$ , corresponding to the locations of innovation or research ( $r$ ), production ( $o$  for origin), and consumption ( $d$  for destination). The two sectors – ICT and non-ICT – are indexed by  $k$  and  $s$ , and economic activity occurs in two types: goods production ( $G$ ) and research ( $R$ ). Time is continuous:  $t^*$  denotes the period in which ideas are produced through research, and  $t$  denotes the period in which goods are produced and consumed. The equilibrium objects and central mechanisms of the model are presented as lemmas, propositions, and corollaries, with proofs provided in Appendix C.

---

<sup>10</sup>By directed innovation I refer to the concept of directed technological change introduced by Acemoglu (2002). I extend this work by introducing the *spatial* direction of technological change.

### 3.1 Microfounded Innovation and Technology Diffusion

In each region and sector, there is a unit continuum of immobile firms. At each time  $t^*$ , these firms hire local innovation workers to conduct research, discovering ideas from a stochastic process with rate  $\lambda_{r,t^*}^k dt^{*11}$ . To capture the ICT shock and the mechanisms through which it shapes the geography of U.S. innovation, I assume that  $\lambda_{r,t^*}^k$  depends on the region's college ratio, sector-specific national research productivity (which incorporates a ***one-time increase in ICT research productivity nationwide, the first component of the ICT shock***), the benefits of colocation with production, and the local supply of inventors. Importantly, all subsequent results in the model hold for any arbitrary specification of  $\lambda_{r,t^*}^k$ .

Each idea pertains to the production of a specific good  $\nu$  on the unit interval  $[0,1]$  and has stochastic quality determined on discovery, such that the aggregate number of ideas discovered in region  $r$  with quality above  $q$  by time  $t$  is drawn from a Poisson distribution with mean parameter<sup>12</sup>:

$$\lambda_{r,t}^Q(q) = q^{-\theta} \int_{-\infty}^t \lambda_{r,t^*} dt^*. \quad (3)$$

After an idea is discovered in region  $r$  at time  $t^*$ , applications of the idea arrive via independent stochastic processes in every region<sup>13</sup>. Each application has stochastic applicability determined on arrival, such that the aggregate number of applications of this idea in region  $o$  at time  $t$  with applicability above  $a$  is drawn from a separate Poisson distribution with mean parameter<sup>14</sup>:

$$\lambda_{ro,t^*t}^A(a) = a^{-\sigma} \cdot \Omega_{ro,t^*}(t - t^*) \cdot \Gamma \left( 1 - \frac{\theta}{\sigma} \right)^{-\frac{\sigma}{\theta}}. \quad (4)$$

From this distribution, the expected number of applications of the idea above  $a$  to ever arrive in region  $o$  is  $\Gamma \left( 1 - \frac{\theta}{\sigma} \right)^{-\frac{\sigma}{\theta}} a^{-\sigma}$ .  $\Omega_{ro,t^*}(t - t^*)^{15}$  is the share of applications that have arrived in region  $o$  by time  $t$  and represents any function of the idea's age  $t - t^*$  with  $\Omega : [0, \infty) \rightarrow [0, 1]$  and  $\lim_{t \rightarrow t^*} \Omega_{ro,t^*}(t - t^*) = 0$ ,  $\lim_{t \rightarrow \infty} \Omega_{ro,t^*}(t - t^*) = 1$ . Intuitively,  $\frac{d\Omega_{ro,t^*}(t - t^*)}{dt}$  captures the density or fertility of the idea in creating new applications. I assume that idea fertility declines exponentially with idea age, following Eaton and Kortum (1999), as formalized below:

---

<sup>11</sup>  $\lambda_{r,t^*}^k$  may be thought of as the output of the research production function

<sup>12</sup> Note that I omit the sector superscript  $k$  in this equation and what follows since the microfounded structure of innovation and technology diffusion is independent across sectors.

<sup>13</sup> This arrival of applications captures technology diffusion at the level of individual ideas.

<sup>14</sup> I provide microfoundations for the Poisson distributions of idea quality and applicability in Appendix C.1.

<sup>15</sup> Notice that  $\Omega_{ro,t^*}(t - t^*)$  could be parameterized as a function of trade and migration flows at time  $t$ , mimicking Buera and Oberfield (2020) and Cai et al. (2025). Because innovation and technology diffusion are fundamentally connected at the level of individual ideas in my model, for any *arbitrary* function  $\Omega_{ro,t^*}(t - t^*)$ , the technology diffusion network then influences innovation rates through the expected value of individual ideas, as shown in Section 3.3 below. This impact cannot be obtained when innovation and technology diffusion are modeled as separate processes.

**Assumption 1.** Idea diffusion occurs exponentially over the idea's age, such that:

$$\Omega_{ro,t^*}(t - t^*) = 1 - e^{-\delta_{ro,t^*}(t-t^*)}, \quad (5)$$

with  $\Omega_{rr,t^*}(t - t^*) = 1$  (i.e. by setting  $e^{-\delta_{rr,t^*}(t-t^*)}$  as 0 for all  $t \geq t^*$ ).

This assumption strikes a balance between two extremes: instantaneous diffusion, where new ideas are disproportionately weighted, and linear diffusion, where older ideas continue to generate numerous new applications over time. In particular, the speed of exponential diffusion is governed by the diffusion lags  $\delta_{ro,t^*}$ . A decrease in the common component  $\delta_{t^*}$  across all region-pairs over time captures **reduced communication costs, the second component of the ICT shock**.

### 3.2 Evolution of Technology Levels, Production, and Trade

The discovery and diffusion of individual ideas generate aggregate trade and technology adoption flows across all region-pairs.

#### 3.2.1 Production of Individual Goods

In each region  $o$ , local firms produce each individual good  $\nu$  using only labor and the most efficient idea application for that good. The efficiency of applying idea  $i$  to good  $\nu$  is the product of its quality and its applicability in that region. Accordingly, the productivity of producing good  $\nu$  in region  $o$  at time  $t$  is:

$$z_{o,t}(\nu) = \max_i \{q_i \cdot a_{i,o,t}\} \quad (6)$$

where  $q_i$  denotes the quality of idea  $i$ , and  $a_{i,o,t}$  denotes its best applicability in region  $o$  at time  $t$ .

Given the stochastic microfoundations of innovation and diffusion detailed in equations (3)-(4), the productivity of each individual good  $z_{o,t}(\nu)$  is drawn from a multivariate Fréchet distribution. Formally:

**Lemma 1.** Given that each idea is discovered at time  $t^*$  at a unique discovery location, idea applications have quality and applicability drawn from the Poisson distributions in equations (3)-(4) respectively, and their efficiency is multiplicative in quality and applicability [equation 6], the **joint productivity distribution** of individual goods across regions at each time  $t$  is multivariate Fréchet and given by:

$$\mathbb{P}[Z_{1,t} \leq z_1, \dots, Z_{N,t} \leq z_N] = \exp \left[ - \sum_{o=1}^N \int_{-\infty}^t \left[ \sum_{r=1}^N \Omega_{ro,t^*}(t - t^*) z_o^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r,t^*} dt^* \right] \quad (7)$$

where  $\rho = 1 - \frac{\theta}{\sigma} < 1$ , and the **marginal productivity distribution** in each region is Fréchet and given by:

$$\mathbb{P}[Z_{o,t} \leq z_o] = \exp \left[ -T_{o,t} z_o^{-\theta} \right] \quad (8)$$

with shape parameter  $\theta > 0$  and scale parameter:

$$T_{o,t} = \sum_{r=1}^N T_{ro,t} = \sum_{r=1}^N \int_{-\infty}^t \underbrace{\Omega_{ro,t^*}(t - t^*)^{1-\rho}}_{\text{technology diffusion}} \cdot \underbrace{\lambda_{r,t^*}}_{\text{innovation}} dt^*. \quad (9)$$

Notice from equation (8) that the marginal productivity distribution in each region is Fréchet, like in the canonical Eaton and Kortum (2002) model [henceforth, EK]. As shown below, this distribution facilitates closed form expressions for trade shares, enhancing tractability. The region's mean productivity in producing individual goods is given by the scale parameter,  $T_{o,t}$ . Unlike EK, this scale parameter is endogenously determined from the *dynamics* of innovation and technology diffusion, as shown by equation (9). Thus, changes in innovation and technology diffusion patterns from the ICT shock – or any other shock to economic fundamentals – impacts trade shares and, consequently, production worker wages and aggregate welfare. Because technology diffusion is modeled as different applications of the same idea, the impact of technology diffusion on trade shares is governed by the relative variability of idea applicabilities against idea quality,  $\rho = \frac{\theta}{\sigma}$ . This elasticity is captured by the exponent  $1 - \rho$  in the marginal (equation 8) and joint productivity distributions (equation 7), ensuring that technology diffusion does not lead to large fluctuations in technology levels and, hence, trade shares over time.

### 3.2.2 Aggregation across Individual Goods and Equilibrium Trade Shares

In each region and sector, a representative final goods producer aggregates across these individual goods  $\nu$  via a Cobb-Douglas function<sup>16</sup>:

$$Y_{d,t} = \exp \int_0^1 \ln Y_{d,t}(\nu) d\nu. \quad (10)$$

Each individual good  $\nu$  need not be produced in the same region  $d$ . Instead, the final goods producer purchases each good from its cheapest production location, taking into account trade costs. Trade costs are of the standard iceberg type, such that delivering one unit of any good from region  $o$  to region  $d$  at time  $t$  requires shipping  $\tau_{od,t} \geq 1$  units of the good, with  $\tau_{oo,t} = 1$  for all  $o$ , and  $\tau_{od,t} \leq \tau_{od',t} \tau_{d'd,t}$  for all  $o$ ,  $d$ , and  $d'$ . Thus, the cost of each good  $\nu$  in region  $d$  is the minimum unit cost across all regions:

$$c_{d,t}(\nu) = \min_o \left\{ \frac{\tau_{od,t} w_{o,t}}{z_{o,t}(\nu)} \right\} \quad (11)$$

where  $w_{o,t}$  is the wage of production workers in region  $o$  at time  $t$ .

Consequently, the equilibrium trade shares are given by the share of goods in each destination that are purchased from different production locations. Because the productivity of each good is drawn from a

---

<sup>16</sup>The Cobb-Douglas aggregator implies that every good is essential for the production of the local sectoral final good, so the monopoly price for each individual good is infinite. Thus, under Bertrand competition introduced in the next subsection, the price charged for every good is a markup on the second lowest cost, simplifying the computation for the expected value of an individual idea. With a CES aggregator, the price of each good is the minimum between the second lowest cost and the monopoly markup of the lowest cost, as shown in Bernard et al. (2003).

multivariate Fréchet distribution, as shown in equation (7), these trade shares can be expressed in closed form. Formally:

**Lemma 2.** *Given that the productivity of each good is drawn from the multivariate Fréchet distribution in equation (7), equilibrium **trade shares** are given by:*

$$\pi_{od,t} = \sum_{r=1}^N \pi_{rod,t} = \sum_{r=1}^N \int_{-\infty}^t \underbrace{\phi_{rd,t^*t}}_{\substack{\text{share of goods in } d \text{ at } t \\ \text{using ideas from } r \text{ at } t^* \\ (\text{idea market shares})}} \cdot \underbrace{\varphi_{o|rd,t^*t}}_{\substack{\text{share of goods in } d \\ \text{produced in } o \text{ at } t \\ \text{given ideas from } r \text{ at } t^* \\ (\text{conditional idea adoption shares})}} dt^* \quad (12)$$

with the **idea market shares** given by:

$$\phi_{rd,t^*t} = \frac{\Phi_{rd,t^*t}^{1-\rho} \lambda_{r,t^*}}{\sum_{r'} \int_{-\infty}^t \Phi_{r'd,t'}^{1-\rho} \lambda_{r',t'} dt'}, \quad (13)$$

where I define the **idea market access** term  $\Phi_{rd,t^*t}$  as:

$$\Phi_{rd,t^*t} \equiv \sum_o \Omega_{ro,t^*} (t - t^*) (w_{o,t} \tau_{od,t})^{-\frac{\theta}{1-\rho}}, \quad (14)$$

and the **conditional idea adoption shares** given by:

$$\varphi_{o|rd,t^*t} = \frac{\Omega_{ro,t^*} (t - t^*) (w_{o,t} \tau_{od,t})^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N \Omega_{ro',t^*} (t - t^*) (w_{o',t} \tau_{o'd,t})^{-\frac{\theta}{1-\rho}}}. \quad (15)$$

The trade shares in equation (12) illustrate the interconnections between innovation, diffusion, and trade<sup>17</sup>. Unlike standard trade models, this model allows for innovation and production to occur in different locations. Consequently, trade shares are represented as the sum of trilateral shares,  $\pi_{rod,t}$ , across all innovation regions  $r$ . These trilateral trade shares denote the share of goods sold in destination region  $d$  that were produced in origin region  $o$  using ideas discovered in region  $r$ . Because the discovery time of an idea,  $t^*$ , may differ from the time of production, trade, and consumption,  $t$ , and because ideas are not perfect substitutes, it is necessary to distinguish goods produced from ideas developed at different times. Accordingly, the trilateral trade shares at time  $t$  are expressed as an integral over all idea cohorts  $t^* \leq t$ :  $\pi_{rod,t} = \int_{-\infty}^t \pi_{rod,t^*t} dt^*$ . The trilateral share for each idea cohort,  $\pi_{rod,t^*t}$ , exhibits a nested structure because productivity is drawn from a multivariate Fréchet distribution with correlation across idea discovery locations. Specifically, the trilateral share for each idea cohort equals the share of goods  $\phi_{rd,t^*t}$  sold in destination region  $d$  at time  $t$  using ideas developed in region  $r$  at time  $t^*$  (idea market shares) multiplied by the conditional probability  $\varphi_{o|rd,t^*t}$  that

---

<sup>17</sup>Note that the trade shares depend only on the probabilistic assumptions of technology (i.e. the Poisson distributions for idea quality and applicability in equations 3 and 4 respectively) and are independent of market structure. This is because each destination always sources goods from their cheapest location, a fundamental result highlighting the flexibility of the Eaton-Kortum world.

region  $o$  is the lowest-cost location for producing these goods (conditional idea adoption shares).

### 3.2.3 Degree of Colocation between Innovation and Production

The idea market and idea adoption shares, in turn, capture the mechanisms that shape the geography of U.S. innovation documented in my empirical findings. Specifically, by aggregating the conditional idea adoption shares in equation (15) across different destination markets  $d$ , we obtain the (unconditional) idea adoption shares:

$$\varphi_{ro,t^*t} = \sum_d \pi_{o|rd,t^*t} = \sum_d \frac{\Omega_{ro,t^*}(t - t^*) (w_{o,t}\tau_{od,t})^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N \Omega_{ro',t^*}(t - t^*) (w_{o',t}\tau_{o'd,t})^{-\frac{\theta}{1-\rho}}}. \quad (16)$$

This expression represents the share of successful ideas discovered in region  $r$  at time  $t^*$  that were adopted in region  $o$  for production of the corresponding goods at time  $t$ . A “successful idea” here refers to one that resulted in a good sold in at least one destination market at time  $t$ . Using this expression, I formally characterize the degree of *colocation* between innovation and production in the following corollary:

**Corollary 1.** *For goods sold at time  $t$  using ideas discovered at time  $t^*$ , the **degree of colocation between innovation and production** for any region  $r$  relative to an alternative production location  $o \neq r$  is given by:*

$$\frac{\varphi_{rr,t^*t}}{\varphi_{ro,t^*t}} = \frac{1}{\Omega_{ro,t^*}(t - t^*)} \cdot \left( \frac{w_{r,t}}{w_{o,t}} \right)^{-\frac{\theta}{1-\rho}} \cdot \frac{\sum_d (\tau_{rd,t})^{-\frac{\theta}{1-\rho}}}{\sum_d (\tau_{od,t})^{-\frac{\theta}{1-\rho}}}. \quad (17)$$

This ratio captures the relative share of goods produced in the idea’s discovery region  $r$  versus an alternative production region  $o$ , providing a direct measure of how closely innovation and production are colocated. The different components reflect key factors driving colocation, including the share of all ideas discovered in region  $r$  at time  $t^*$  that have diffused to  $o$  (first term), the cost competitiveness between  $r$  and  $o$  (second term), and the relative accessibility of destination markets  $d$  from  $r$  and  $o$  (third term). In my application to U.S. innovation, this ratio provides two reasons why ICT production became more concentrated in high-skill regions following a one-time increase in ICT research productivity (the first exogenous component of the ICT shock): frictions in technology diffusion (the first term), and the greater accessibility of ICT destination markets from high-skill regions.

### 3.2.4 The Asymmetric Scale Effect of Rising Technology Diffusion Speeds

Additionally, the idea market shares (equation 13) give rise to what I term the *asymmetric scale effect*: high-skill regions experience a larger increase in idea market access (equation 14) from a uniform rise in bilateral diffusion speeds. Substituting Assumption 1 into equation 14 yields the following corollary:

**Corollary 2.** *If bilateral diffusion speeds are symmetric ( $\delta_{rr',t^*} = \delta_{r'r,t^*} \equiv \delta_{rr'}$ ) and trade costs are identical across all region pairs ( $\tau_{rd,t} = \tau_{r'd,t} \equiv \tau_t$ ), then an increase in the bilateral diffusion speed raises idea market*

access more for the higher-wage region, i.e.:

$$\frac{\partial \Phi_{rd,t^*t}}{\partial \delta_{rr',t^*}} - \frac{\partial \Phi_{r'd,t^*t}}{\partial \delta_{rr',t^*}} = \delta_{rr',t^*} e^{-\delta_{rr',t^*}(t-t^*)} \left[ (w_{r',t}\tau_{r'd,t})^{-\theta} - (w_{r,t}\tau_{rd,t})^{-\theta} \right] > 0 \quad \text{if } w_{r,t} > w_{r',t}.$$

Intuitively, faster diffusion raises idea market access more in high-wage (and high-skill) regions because it allows them to exploit the lower production costs of other regions when applying their ideas. By contrast, the low-wage regions gain less: their ideas already benefit from relatively low local production costs, so improved access to higher cost regions for production yields only modest additional gains. This asymmetric scale effect helps explain why firms with R&D in high-skill CZs expanded geographically far more than firms with R&D in low-skill CZs, as documented in Section 2.4.2.

Corollaries 1 and 2 show how the components of the trade shares generate two key endogenous mechanisms: the colocation of ICT innovation and production in high-skill regions, and the asymmetric scale effect. These mechanisms arise only in a model where innovation and technology diffusion are connected at the level of individual ideas<sup>18</sup>. To enable these mechanisms to shape the trajectory of U.S. innovation geography following the ICT shock, I now *introduce directed innovation in a dynamic spatial model* through imperfect competition and dynamic worker mobility. Imperfect competition generates profits from individual ideas that shape the incentives of workers, while their mobility determines the evolution of innovation levels in each region and sector.

### 3.3 Market Structure, Prices, Profits, and Innovation Worker Wages

In each region and sector, firms hire innovation workers and own their ideas. These firms engage in Bertrand competition, where the lowest cost producer of each good claims the entire market for that good, charging the highest markup that deters any competitor from entering.

#### 3.3.1 Prices and the Markup Distribution

Since all goods are essential in the production of the local final good (equation 10), the price of each good is its second lowest cost. Aggregating across all goods yields the price index in each region. Formally:

**Lemma 3.** *Given Bertrand competition and the Cobb-Douglas aggregation of goods in each region and sector (equation 10, the **price index** in region  $d$  at time  $t$  is:*

---

<sup>18</sup>When innovation and technology diffusion are modeled at an aggregate level (e.g. Desmet et al., 2018) or as independent processes (e.g. Buera and Oberfield, 2020; Cai et al., 2025), the idea *adoption* shares would exactly coincide with the idea *diffusion* shares, so the degree of colocation between innovation and production does not depend on wages and trade costs. Additionally, idea market access would remain unchanged when the speed of technology diffusion changes, so the asymmetric scale effect cannot be obtained. When innovation and technology diffusion are connected at the level of individual ideas but ideas have identical efficiency across different production locations (Cai et al., 2022; Eaton and Kortum, 2024), colocation and the asymmetric effect can be obtained, but technology diffusion results in large fluctuations in idea market shares and hence trade shares (as seen by setting  $\rho = 0$  in equation 13).

$$P_{d,t} = \gamma \left[ \sum_{r'=1}^N \int_{-\infty}^t \Phi_{r'd,t^*t}^{1-\rho} \lambda_{r',t^*} dt^* \right]^{-\frac{1}{\theta}}, \quad (18)$$

where  $\gamma$  is the Euler-Mascheroni constant.

Additionally, the markup for each good is the ratio between the second lowest cost (from the second highest quality draw) and the lowest cost (from the highest quality draw). Since idea quality is drawn from a Pareto distribution, the ratio of the second highest quality to the highest quality draw conditional on the second highest quality draw is always Pareto and scale invariant. Consequently, the distribution in which markups are drawn is Pareto and invariant across regions and time. Formally:

**Lemma 4.** *Given Bertrand competition and the Cobb-Douglas aggregation of goods in each region and sector (equation 10), the **markup distribution** is Pareto and invariant across idea discovery time  $t^*$ , production time  $t$ , idea discovery location  $r$ , origin  $o$ , destination  $d$ , and sector  $k$ :*

$$G^{(2)/(1)}(m) \equiv \mathbb{P} \left[ \frac{C^{(2)}}{C^{(1)}} \leq m | C^{(2)} = c_2 \right] = 1 - m^{-\theta}. \quad (19)$$

This distribution coincides with that in a setting without technology diffusion because I assume that only the most efficient application of each idea can be used to produce its corresponding good (see equation 6). Consequently, Bertrand competition occurs across ideas, not across multiple applications of the same idea.<sup>19</sup> This invariant markup distribution is crucial for deriving closed-form expressions for aggregate profits and the expected value of individual ideas within each region-sector.

### 3.3.2 Profits and Innovation Worker Wages

Specifically, the invariant markup distribution implies that all firms in sector  $k$  selling in destination  $d$  charge a markup drawn from  $G^{(2)/(1)}(m)$ <sup>20</sup>. Consequently, the total profits earned from selling in destination  $d$ , irrespective of the locations of innovation and production, are given by:

$$\Pi_{d,t} = X_{d,t} \int_0^1 1 - \frac{1}{m(\nu)} d\nu = X_{d,t} \int_1^\infty 1 - \frac{1}{m(\nu)} dG^{(2)/(1)}(m) = \frac{X_{d,t}}{1 + \theta}, \quad (20)$$

where  $X_{d,t}$  is the total spending by production workers and inventors in region  $d$  at time  $t$  and given by equation (24). Because the markup distribution is identical whether conditional or unconditional on the production location, profits earned by firms from selling their goods can be arbitrarily assigned to production,

---

<sup>19</sup>I impose this simplifying assumption because I interpret technology diffusion in my model as primarily occurring *within* firms, to capture the effects of firm spatial expansion following the ICT shock. This interpretation is also consistent with the strong intellectual property protections in the US. Allowing for both within- and across-firm diffusion, and hence competition across different applications of the same idea, may be more appropriate in an international or developing-country context.

<sup>20</sup>Formally, to derive the unconditional markup distribution across all ideas, we need to integrate the conditional markup distribution  $G^{(2)/(1)}(m)$  over all values of the second lowest cost  $c_2$ . However, since this conditional distribution is independent of  $c_2$ , it is exactly equal to the unconditional markup distribution.

innovation, or a combination of both<sup>21</sup>. I now make assumptions on how these profits are allocated and subsequently paid to individual inventors.

**Assumption 2** (Allocation of Profits from Sales).

- (i) *There is perfect intellectual property protection, such that all profits from sales are allocated to innovation;*
- (ii) *At every instant, firms invest their profits in risk-free assets, which are produced using the same technology as final goods in each region<sup>22</sup> and do not depreciate over time. Firms compensate the inventors they hire with wages equal to the expected return of their innovation efforts.*

Given Assumption 2(i), the expected value of an idea is formalized in the following lemma:

**Lemma 5.** *Given Bertrand competition, the Cobb-Douglas aggregation across individual goods (equation 10), the time- and region-invariant markup distribution (equation 19), and the allocation of profits from sales (Assumption 2(i)), the **expected value of an idea** in region  $r$  and sector  $k$  is:*

$$\check{V}_{r,t^*} = \int_{t^*}^{\infty} e^{-\zeta(t-t^*)} \sum_{d=1}^N \underbrace{\frac{\phi_{rd,t^*}}{\lambda_{r,t^*}}}_{\substack{\text{share of profits earned} \\ \text{in region } d \text{ at time } t \\ \text{by an idea discovered} \\ \text{in region } r \text{ at time } t^*}} \cdot \underbrace{\frac{X_{d,t}}{1+\theta}}_{\substack{\text{profits earned} \\ \text{in region } d \\ \text{at time } t \\ \text{by all ideas}}} \cdot \underbrace{\frac{P_{rt^*}}{P_{rt}}}_{\substack{\text{accounting for} \\ \text{changes in} \\ \text{purchasing power} \\ \text{over time}}} dt \quad (21)$$

where  $\theta$  is the trade elasticity,  $\zeta$  is the discount rate, and  $\phi_{rd,t^*}$  is the idea market share.

The expected value of an idea is given by the present discounted value of the trajectory of profits for all  $t \geq t^*$  during which the idea is used. The first term captures the probability the idea developed in region  $r$  at time  $t^*$  results in a good sold in destination market  $d$  at time  $t$ , the second term represents the total profits earned in destination market  $d$  at time  $t$  by all ideas, while the third term accounts for changes in purchasing power over time.

Given Assumption 2(ii), wages of inventors are simply the product of the expected number of ideas they discover, as described by the idea production function in equation (30), and the expected value of each idea in equation (5). Formally:

$$w_{r,t^*}^{k,R} = \frac{\lambda_{r,t^*}^k}{L_{r,t^*}^{k,R}} \check{V}_{r,t^*}^k. \quad (22)$$

### 3.4 Consumption, Market Clearing, and Production Worker Wages

At each time  $t$ , production workers and inventors have identical Cobb-Douglas preferences over local final goods from the ICT and non-ICT sectors, with  $\iota$  capturing the sectoral expenditure share allocated to goods produced in the ICT sector. Given these preferences, the market clearing condition at each time  $t$  is

---

<sup>21</sup>Consequently, my model is consistent with any structure of intellectual property protection or technology licensing, nesting Hemous et al. (2025)

<sup>22</sup>This structure mirrors the investment good technology in Kleinman et al. (2023), ensuring that the market clearing condition in each period (see equation 23) is solely determined by contemporaneous variables.

formalized in the following lemma:

**Lemma 6.** *Given the allocation of profits from sales in Assumption 2, the combined goods and innovation market clearing condition at time  $t$  is given by:*

$$\frac{1+\theta}{\theta} w_{o,t}^k L_{o,t}^k = \sum_d \pi_{od,t}^k \iota^k \left[ \sum_s \left( w_{d,t}^s L_{d,t}^s + \sum_r \varphi_{dr,t}^s \frac{1}{\theta} w_{r,t}^s L_{r,t}^s \right) \right] \quad (23)$$

where  $\pi_{od,t}^k$  are the trade shares from equation (12),  $\varphi_{dr,t}^s$  are the idea adoption shares from equation (16),  $w_{o,t}^k$  are the production worker wages, and  $L_{o,t}^k$  is the number of production workers.

In particular, total expenditure in region  $d$  on sector  $k$  goods is given by:

$$X_{d,t}^k = \iota^k \left[ \sum_s \left( w_{d,t}^s L_{d,t}^s + \sum_r \varphi_{dr,t}^s \frac{1}{\theta} w_{r,t}^s L_{r,t}^s \right) \right]. \quad (24)$$

Since profits earned from production are a constant multiple of the income earned by production workers in the region-sector, and firms reinvest their profits in the same period to produce assets, neither profits nor wages of innovation workers appear explicitly in the market clearing condition. Consequently, I omitted the superscript  $G$  for wages and labor to reduce notational burden.

### 3.5 Dynamic Worker Mobility

To allow innovation and production wages (equations 22 and 23 respectively) to shape the spatial and sectoral directions of innovation, I introduce worker mobility across regions and sectors and between production and research. I model mobility decisions as dynamic to capture how anticipated future opportunities influenced worker sorting to high-skill CZs in the 1980s. These dynamic bilateral decisions require additional notation. I denote worker occupations  $h, n$  to represent either production (G) or research (R). Because continuous-time mobility complicates the separation of current and future payoffs, I impose the following assumption on the timing of migration:

**Assumption 3.** *A Poisson arrival process with rate 1 governs when all workers can move.*

In anticipation of when a move arrival occurs at some  $t'$ , workers decide where to migrate to, which sector to work in (ICT or non-ICT), and whether to engage in production or research based on the present value stream of utility minus the associated mobility costs from their current region, sector, and occupation. With perfect foresight, the optimization problem of a worker in region  $d$ , sector  $k$ , and occupation  $h$  at time  $t$  is:

$$v_{d,t}^{k,h} = \max_{o,s,n} \mathbb{E}_t \left( \int_t^{t'} \frac{w_{d,\tilde{t}}^{k,h}}{P_{d,\tilde{t}}} d\tilde{t} \right) + \frac{1}{1+\zeta} \mathbb{E}_t \left( \mathbb{E}_\epsilon \left[ v_{o,t'}^{s,n} \right] \right) - \kappa_{do,t}^{ks,hn} + \epsilon_{o,t}^{s,n} \quad (25)$$

where  $\kappa_{do,t}^{ks,hn}$  are the costs of moving from region  $d$ , sector  $k$ , and occupation  $h$  to region  $o$ , sector  $s$ , and occupation  $n$ ,  $\zeta$  is the discount rate,  $\epsilon_{o,t}^{s,n}$  is an individual-specific idiosyncratic shock in each potential

destination region-sector-occupation,  $\mathbb{E}_t(\cdot)$  is the time- $t$  expectation over future state variables, and  $\mathbb{E}_\epsilon(\cdot)$  is the expectation over the agent's future realizations of the idiosyncratic shock. At each time  $t$ , I assume each individual-specific idiosyncratic shock  $\epsilon_{o,t}^{s,n}$  is drawn from a multivariate Gumbel distribution with the following cumulative distribution function:

$$\check{F}\left(\{\epsilon_{o,t}^{s,n}\}_{o=1,\dots,N}^{s=\{\text{ICT,non-ICT}\},n=\{G,R\}}\right) = \exp\left\{-\left[\sum_o \sum_s \left(\sum_n \exp(-\epsilon_{o,t}^{s,n})^{\frac{\gamma}{v}}\right)^v\right]\right\} \quad (26)$$

where  $v$  is the elasticity of worker mobility across regions and sectors, and  $\frac{\gamma}{v}$  is the elasticity of worker mobility between production and research.

**Lemma 7.** *Given individual-level worker mobility decisions defined by equations (25)-(26), the expected value or lifetime utility of a representative worker in labor market  $(d, k, h)$  is given by:*

$$V_{d,t}^{k,h} \equiv \mathbb{E}_\epsilon\left[v_{d,t}^{k,h}\right] = \mathbb{E}_t\left(\int_t^{t'} \frac{w_{d,\tilde{t}}^{k,h}}{P_{d,\tilde{t}}} d\tilde{t}\right) + \frac{1}{\gamma} \log \left[ \sum_o \sum_s \left( \sum_n \exp\left(\frac{1}{1+\zeta} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn}\right)^{\frac{\gamma}{v}} \right)^v \right] \quad (27)$$

aggregate **mobility shares** of workers from  $(d, k, h)$  to  $(o, s, n)$  is given by:

$$\begin{aligned} \mu_{do,t}^{ks,hn} &\equiv \mathbb{E}_t\left[\tilde{\mu}_{do,t'}^{ks,hn}\right] \equiv \mu_{do,t}^{ks,hn} | \mu_{do,t}^{ks} \cdot \mu_{do,t}^{ks} \\ &= \frac{\exp\left(\frac{1}{1+\zeta} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn}\right)^{\frac{\gamma}{v}}}{\underbrace{\sum_{n'} \exp\left(\frac{1}{1+\zeta} V_{o,t'}^{s,n'} - \kappa_{do,t}^{ks,hn'}\right)^{\frac{\gamma}{v}}}_{\text{switching between production and research}}} \cdot \frac{\left[\sum_{n'} \exp\left(\frac{1}{1+\zeta} V_{o,t'}^{s,n'} - \kappa_{do,t}^{ks,hn'}\right)^{\frac{\gamma}{v}}\right]^v}{\underbrace{\sum_{o'} \sum_{s'} \left[\sum_{n'} \exp\left(\frac{1}{1+\zeta} V_{o',t'}^{s',n'} - \kappa_{do',t}^{ks',hn'}\right)^{\frac{\gamma}{v}}\right]^v}_{\text{migration across regions and sectors}}} \end{aligned} \quad (28)$$

and the worker population in  $(o, s, n)$  evolves as follows:

$$L_{o,t'}^{s,n} = \sum_h \sum_k \sum_d \mu_{do,t}^{ks,hn} L_{d,t}^{k,h}. \quad (29)$$

### 3.6 Endogenous and Directed Innovation

Up to this point, the Poisson arrival rate of new ideas (or innovation levels)  $\lambda_{r,t^*}^k$  has been treated as exogenous. To introduce endogenous and directed innovation, I allow  $\lambda_{r,t^*}^k$  depend on the number of innovation workers in the respective region and sector (third term in equation 30), so that higher innovation employment increases innovation output. As profits and innovation wages rise in a given region, additional workers are drawn into research here, endogenously increasing local innovation. To connect the model to single-region endogenous growth models such as Romer (1990), I assume  $\lambda_{r,t^*}^k$  is linear in the region-sector technology level  $T_{r,t^*}^k$ <sup>23</sup> (fifth term in equation 30). Finally, to align my model with the mechanisms through which the

<sup>23</sup> Alternatively, replacing  $T_{l,t^*}$  with  $T_{l,t^*}^\beta$  with  $\beta < 1$  would yield semi-endogenous growth, as in many single-region growth models pioneered by Jones (1995); Kortum (1997).

ICT shock influenced the geography of U.S. innovation after 1990 – documented in the empirical section – I allow  $\lambda_{r,t^*}^k$  to depend on additional components.

Formally,  $\lambda_{r,t^*}^k$  depends on five terms:

$$\lambda_{r,t^*}^k = \underbrace{A_{r,t^*}}_{\text{fundamental research productivity in region } r \text{ (function of college ratio)}} \cdot \underbrace{A_{t^*}^k}_{\text{sector-specific national research productivity (Direct effect of ICT)}} \cdot \underbrace{T_{r,t^*}^k}_{\text{sector-specific technology level in region } r} \cdot \underbrace{\left(L_{r,t^*}^{k,G}\right)^\chi}_{\text{number of production workers (benefits of colocation with production)}} \cdot \underbrace{L_{r,t^*}^{k,R}}_{\text{number of inventors}} \quad (30)$$

The first term represents the region's fundamental research productivity. It includes a time-invariant component proportional to the region's college ratio – capturing why high-skill regions innovate more – and a time-varying component that accounts for factors unexplained by the model, such as local innovation policies, infrastructure, or other regional trends. The second term denotes the sector-specific research productivity. As described at the beginning of my model, this term incorporates the first component of the ICT shock via a *one-time* increase in ICT research productivity  $A_{1990}^{ICT}$ , ensuring that ICT innovation levels rise in all regions relative to a scenario without this shock. The fourth term incorporates the benefits of colocation between innovation and production, modeled as contemporaneous spillovers from the number of production workers in the same region and sector, where  $0 < \chi < 1$ .

With the specification of  $\lambda_{r,t^*}^k$ , I now define the equilibrium of my model as follows:

### 3.7 Definition of Equilibrium

Given an initial distribution of technology levels  $\{T_{o,0}^k\}_{o=1,k=1}^{N,N}$  and workers  $\{L_{o,0}^{k,h}\}_{o=1;k,h}^{N;\{\text{ICT,non-ICT}\};\{G,R\}}$ , trajectories of bilateral trade costs  $\{\tau_{od,t}\}_{o=1,d=1,t=0}^{N,N,\infty}$ , bilateral migration costs  $\{\kappa_{od,t}^{ks,ha}\}_{o=1;d=1;t=0;k,s;h,a}^{N;N;\infty;\{\text{ICT,non-ICT}\};\{G,R\}}$ , bilateral diffusion lags  $\{\delta_{od,t^*}\}_{o=1,d=1,t^*=0}^{N,N,\infty}$ , fundamental productivities in idea production  $\{A_{r,t^*}, A_{t^*}^k\}_{r=1,k,t^*=0}^{N,\{\text{ICT,non-ICT}\},\infty}$  and fundamental parameters and elasticities  $\{\theta, \sigma, \nu, \Upsilon, \iota, \alpha, \zeta\}$ , the **dynamic competitive equilibrium** is defined by a trajectory of values, wages, prices and labor allocations  $\{V, w, P, L\}$  that satisfy the bilateral migration shares and evolution of worker populations (Lemma 7), innovation levels (equation 30), evolution of technology levels (Lemma 1 equation 9), bilateral trade and idea adoption shares (Lemma 2 equations 12 and 16), price indices (Lemma 3), returns to innovation (Lemma 5) and market clearing condition (Lemma 6).

### 3.8 Model Interpretation and Extensions

#### Spatial Direction of Innovation

My model introduces *endogenous and directed innovation in a spatial setting*. To highlight this methodological contribution and connect it to the sectoral direction of technical change developed by Acemoglu (1998, 2002, 2007), I now leverage the expression for inventor wages (equation 22) and my specification of

the research production function (equation 30) to define the spatial direction of innovation in my setting, as summarized in the following proposition:

**Proposition 1.** *The spatial direction of innovation is governed by the ratio of sector-specific inventor real wages across regions, given as follows:*

$$\frac{\omega_{r,t^*}^{k,R}}{\omega_{r',t^*}^{k,R}} = \underbrace{\frac{A_{r,t^*}}{A_{r',t^*}}}_{\text{fundamental research productivity (function of college ratio)}} \cdot \underbrace{\frac{T_{r,t^*}^k}{T_{r',t^*}^k}}_{\text{relative technology levels}} \cdot \underbrace{\left( \frac{L_{r,t^*}^{k,G}}{L_{r',t^*}^{k,G}} \right)^\chi}_{\text{benefits of colocation of innovation and production}} \cdot \underbrace{\frac{\int_{t^*}^{\infty} e^{-\zeta(t-t^*)} \sum_{d=1}^N \frac{\phi_{rd,t^*t} Y_{d,t}}{1+\theta P_{r,t}} dt}{\int_{t^*}^{\infty} e^{-\zeta(t-t^*)} \sum_{d=1}^N \frac{\phi_{r'd,t^*t} Y_{d,t}}{1+\theta P_{r',t}} dt}}_{\text{expected market potential of an idea, with the idea market shares capturing the asymmetric scale effect}} \quad (31)$$

The higher the ratio in equation (31), the greater the returns to innovation directed toward region  $r$  relative to region  $r'$ . Within each sector, four main factors shape the incentives to innovate across space: (i) time-varying fundamental research productivity, including a time-invariant component tied to regional college ratios; (ii) relative technology levels; (iii) the benefits of colocation between innovation and production; and (iv) the expected market potential of an idea discovered in that region and sector. The first three factors arise from the research production function that determines the Poisson arrival rate of ideas (equation 30), while the fourth factor is the expected value of individual ideas (equation 21). Together, these elements capture the endogenous mechanisms that shape the geography of innovation along the transition path, mirroring the empirical patterns documented in Section 2.4. More precisely, any increase in the ratio in equation (31) following the ICT shock – driven by one or more of these mechanisms – induces innovation workers to move from region  $r'$  to region  $r$ , thereby raising innovation in  $r$  relative to  $r'$  through equation (30).

This ratio serves as the spatial analog to the sectoral direction of technical change: it captures inventor real wages in partial equilibrium and thus highlights the key forces shaping where innovation is directed. It is important to emphasize, however, that inventor real wages are jointly determined in general equilibrium together with technology levels, labor allocations, prices, and idea market shares in equation (31). Consequently, inventor real wages – and thus the spatial direction of innovation – are also indirectly influenced by many other factors in the economy, such as production worker wages across all regions. Appendix C.8 presents the corresponding wage ratio that governs the sectoral direction of innovation.

## Model Extensions

My dynamic spatial model microfounds innovation and technology diffusion *within* the Eaton-Kortum framework. Consequently, it is highly tractable and can accommodate a broad range of extensions. These include dynamic worker sorting, input-output loops, capital accumulation, and the inclusion of amenities with congestion and agglomeration in goods production, as demonstrated in Appendix D.

## 4 Aggregate Consequences of the Rising Spatial Concentration of Innovation from the ICT Shock

Apart from introducing *endogenous and directed innovation in a spatial setting* to capture the key mechanisms driving the rising concentration of U.S. innovation in high-skill CZs, my dynamic spatial model provides a flexible framework to *decompose the welfare impact of any shock to economic fundamentals into its transitory and long-run growth components*. In this section, I develop the methodological tools to obtain this decomposition , which illustrates the aggregate consequences of the ICT shock.

### 4.1 Balanced Growth Path

I first show that the balanced growth path is highly flexible due to its block-recursive nature. This structure both distinguishes and connects the *causes* and *consequences* of the geography of innovation, while establishing a direct link between quantitative trade and spatial models with innovation and endogenous growth models in macroeconomics. The following proposition formalizes this structure:

**Proposition 2.** *Along the **balanced growth path** where all equilibrium variables grow at constant (but possibly different) rates:*

(i) *The growth rate of technology in each sector  $g^k$  is identical across regions and determined by the solution to the following system of equations:*

$$\dot{T}_o^k(t) = \sum_r \gamma_r^k T_r^k(t) \int_{-\infty}^t g^k e^{-g^k(t-t^*)} \Omega_{ro}(t-t^*)^{1-\rho} dt^* \quad (32)$$

where  $\gamma_r^k$  is the endogenous region-sector-specific innovation rate. In matrix form, this equation is given by:

$$g\mathbf{T}^k = \Delta^k(g)\mathbf{T}^k \quad (33)$$

where  $\mathbf{T}^k$  is an  $N \times 1$  vector with representative element  $T_o^k$  and  $\Delta^k(g)$  is an  $N \times N$  matrix with representative element:

$$\Delta_{ro}^k(g^k) = \gamma_r^k \int_0^\infty g^k e^{-g^k a} \Omega_{ro}(a)^{1-\rho} da.$$

Thus,  $g^k$  is the Perron-Frobenius root of equation (33) with relative technology levels  $\mathbf{T}$  corresponding to the Perron-Frobenius eigenvector that is defined up to a scalar multiple;

(ii) *Production worker wages, inventor wages, and the distribution of workers across regions, sectors, and occupations are constant, prices are falling at rate  $g_p = \frac{1}{\theta} \sum_k \iota^k g^k$ , and the expected value of workers is rising at rate  $g_v = \frac{1+\zeta}{\zeta} \frac{1}{\theta} \sum_k \iota^k g^k$ .*

On aggregate consequences, part (i) of the proposition shows how any arbitrary steady-state geography of innovation rates ( $\gamma_r^k = \frac{\lambda_r^k}{T_r^k}$ ) and region-pair idea diffusion speeds ( $\delta_{ro}$ ) can generate a balanced growth path characterized by parallel growth at the sectoral rate  $g^k$ , alongside persistent level differences in technology  $T^k$  across regions for each sector<sup>24</sup>. Notably, reduced communication costs resulting from the ICT shock

---

<sup>24</sup>This result arises because the Poisson arrival rate of ideas depends linearly on technology levels (see equation 30)

directly impact aggregate growth, in addition to shaping the geography of innovation. Part (ii) shows how the sector-specific technology growth rates ( $g^k$ ) determine the rate at which prices fall and, consequently, the growth rate of workers' expected value on the balanced growth path. This price decline parallels standard macroeconomic growth models. However, unlike those models, the rate at which prices fall here depends on the entire spatial distribution of innovation, rather than solely on the aggregate innovation rate.

On *causes*, part (ii) of the proposition shows that the steady-state geography of innovation rates ( $\gamma_r^k$ ) is determined by the steady-state distribution of workers, which in turn is shaped by the exogenous fundamentals of the economy, including bilateral idea diffusion speeds, migration costs, trade costs, and research productivities. The equilibrium expressions for innovation rates, inventor wages and other variables are provided in Appendix C.9. In steady-state, this block of the model mirrors a standard quantitative trade model. Since any arbitrary steady-state geography of innovation rates is consistent with the balanced growth path described above, my model can readily accommodate various mechanisms that drive the geography of innovation in different contexts.

## 4.2 Transition Path

To better understand how shocks to the economy's fundamentals shape the steady-state distribution of workers and innovation rates, I now characterize the transition path. Specifically, I assume the economy is initially on a balanced growth path<sup>25</sup> and characterize its transition to a new balanced growth path following shocks to fundamentals at time  $t = 0$ . To align my model with observed annual data on innovation levels, trade shares, migration shares, and firm establishment networks, I strengthen Assumption 3 as follows:

**Assumption 4.** *A Poisson arrival process with rate 1 governs when innovation, production, consumption, and migration occurs across all agents and all regions*<sup>26</sup>.

Accordingly, I now interpret  $\lambda_{r,t^*}^k$  as the total number of ideas discovered in region  $r$  and sector  $k$  during the discrete time period  $t^*$ , rather than the Poisson arrival rate of new ideas at the instant  $t^*$  in continuous time. Given this assumption, I then extend the **dynamic hat algebra** methodology developed by Caliendo et al. (2019) to simulate the transition path without information on levels of migration and trade costs. Specifically, let  $\mathcal{T}$  be the set of times where innovation, production, consumption, and migration occur,  $\tilde{x}_t = x_t e^{-g_x t}$  denote the detrended value with growth rate  $g_x$  and  $\dot{x}_t = \frac{x_{t'}}{x_t}$  represent changes over time for any variable  $x$ , and  $u = \exp(V)$ . The transition path can then be obtained given the trajectory of time changes in migration and trade costs – as opposed to levels – alongside levels of research productivity and diffusion lags. This characterization is formalized in the following proposition:

---

<sup>25</sup>This assumption substitutes for historical data on innovation levels and diffusion speeds from time  $-\infty$  to 0, enabling computation of trade shares at all time periods.

<sup>26</sup>This assumption also ensures that workers' expected values are equalized across regions, sectors, and between production and research when innovation and production occur. An alternative, stronger assumption is that all these activities occur only at the beginning or end of each year. In expectation, both assumptions are equivalent.

**Proposition 3.** Given an initial distribution of workers, **wages**, technology levels, and migration and trade shares  $\left\{L_{o,0}^k, T_{o,0}^k, \mu_{od,0}^{ks,hn}, \pi_{od,0}^k\right\}_{o=1; d=1; k,s; h,n}^{N; N; \{ICT, non-ICT\}; \{G, R\}}$ , exogenous trajectories of changes in migration and trade costs, research productivity, and bilateral diffusion lags  $\left\{\dot{\kappa}_{od,t}^{ks,hn}, \dot{\tau}_{od,t}, A_{r,t}^k, \delta_{ro,t}\right\}_{o=1; d=1; r=1; k,s; h,n; t=0}^{N; N; N; \{ICT, non-ICT\}; \{G, R\}; \infty}$ , the **transition path** of the economy is characterized by the evolution of the distribution of workers for all  $t^*, t, t' \geq 0$ :

$$\log \left( \dot{\tilde{u}}_{d,t}^{k,h} \right) = \log \left( \frac{\dot{w}_{d,t}^{k,h}}{\dot{\tilde{P}}_{d,t}} \right) + \frac{1}{\Upsilon} \log \left[ \sum_n \left[ \sum_s \sum_o \mu_{do,t'}^{ks,hn} \left( \dot{\tilde{u}}_{o,t'}^{s,n} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \dot{\kappa}_{od,t}^{ks,hn} \right)^{\frac{\Upsilon}{v}} \right]^v \right] \quad (34)$$

$$\mu_{od,t'}^{ks,hn} = \frac{\mu_{od,t}^{ks,hn} \left( \dot{\tilde{u}}_{d,t'}^{s,n} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \dot{\kappa}_{od,t}^{ks,hn} \right)^{\frac{\Upsilon}{v}}}{\sum_{n'} \mu_{od,t}^{ks,hn'} \left( \dot{\tilde{u}}_{d,t'}^{s,n'} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \dot{\kappa}_{od,t}^{ks,hn'} \right)^{\frac{\Upsilon}{v}}} \cdot \frac{\left[ \sum_{n'} \mu_{od,t'}^{ks,hn'} \left( \dot{\tilde{u}}_{d,t'}^{s,n'} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \dot{\kappa}_{od,t}^{ks,hn'} \right)^{\frac{\Upsilon}{v}} \right]^v}{\sum_{n'} \left[ \sum_{d'} \sum_{s'} \mu_{od',t'}^{ks',hn'} \left( \dot{\tilde{u}}_{d',t'}^{s',n'} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \dot{\kappa}_{od',t}^{ks',hn'} \right)^{\frac{\Upsilon}{v}} \right]^v} \quad (35)$$

$$L_{d,t'}^{k,h} = \sum_n \sum_s \sum_o \mu_{od,t'}^{ks,hn} L_{o,t}^{s,n} \quad (36)$$

where at each time  $t$  innovation levels are given by equation (30) and technology levels by:

$$\begin{aligned} \lambda_{r,t}^k &= A_{t^*}^k A_{r,t^*} \left( L_{r,t^*}^{k,G} \right)^\chi \left( L_{r,t^*}^R \right)^\alpha L_{r,t^*}^{k,R} T_{r,t^*}^k \\ T_{o,t}^k &= \sum_{r=1}^N \sum_{t^* \in \mathcal{T}_{-\infty}^t} \Omega_{ro,t^*} (t - t^*)^{1-\rho} \cdot \lambda_{r,t^*}^k, \end{aligned} \quad (37)$$

the trade equilibrium is given by:

$$\pi_{od,t}^k = \sum_{r=1}^N \sum_{t^* \in \mathcal{T}_{-\infty}^t} \frac{\Omega_{ro,t^*} (t - t^*) \left( w_{o,t}^{k,G} \tau_{od,t}^k \right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'} \Omega_{ro',t^*} (t - t^*) \left( w_{o',t}^{k,G} \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}}} \frac{\left[ \sum_{o'} \Omega_{ro',t^*} (t - t^*) \left( w_{r,t}^{k,G} \tau_{ro',t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}}{\sum_{r'} \sum_{\check{t} \in \mathcal{T}_{-\infty}^t} \left[ \sum_{o'} \Omega_{r'o',\check{t}} (t - \check{t}) \left( w_{o',t}^{k,G} \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}} \lambda_{r,t^*}^k \quad (38)$$

$$\varphi_{ro,t}^k = \sum_{d=1}^N \sum_{t^* \in \mathcal{T}_{-\infty}^t} \frac{\Omega_{ro,t^*} (t - t^*) \left( w_{o,t}^{k,G} \tau_{od,t}^k \right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'} \Omega_{ro',t^*} (t - t^*) \left( w_{o',t}^{k,G} \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}}} \frac{\left[ \sum_{o'} \Omega_{ro',t^*} (t - t^*) \left( w_{r,t}^{k,G} \tau_{ro',t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}}{\sum_{r'} \sum_{\check{t} \in \mathcal{T}_{-\infty}^t} \left[ \sum_{o'} \Omega_{r'o',\check{t}} (t - \check{t}) \left( w_{o',t}^{k,G} \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}} \lambda_{r',\check{t}}^k \quad (39)$$

$$P_{d,t}^k = \Gamma \left[ \sum_{r'=1}^N \sum_{t^* \in \mathcal{T}_{-\infty}^t} \left[ \sum_{o'=1}^N \Omega_{r'o',t^*} (t - t^*) \left( w_{o',t}^{k,G} \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r',t^*}^k \right]^{-\frac{1}{\theta}} \quad (40)$$

$$\frac{1+\theta}{\theta} w_{o,t}^{k,G} L_{o,t}^{k,G} = \sum_d \pi_{od,t}^k \left[ \sum_k \left( w_{d,t}^{k,G} L_{d,t}^{k,G} + \sum_r \varphi_{dr,t}^k \frac{1+\theta}{\theta} w_{r,t}^{k,G} L_{r,t}^{k,G} \right) \right]$$

where the market clearing condition comes from equation (23), and the wages of inventors, returns to inno-

vation and the probability that goods sold in destination  $d$  at time  $t'$  uses ideas discovered in region  $r$  at time  $t$  are given by:

$$w_{r,t}^{k,R} = \frac{\check{V}_{r,t}^k \lambda_{r,t}^k}{L_{r,t}^{k,R}} \quad (41)$$

$$\check{V}_{r,t}^k = \sum_{t' \in \mathcal{T}_t^\infty} \left( \frac{1}{1-\rho} \right)^{t'-t} \sum_d \frac{X_{d,t'}^k}{1+\theta} \cdot \frac{P_{r,t}}{P_{r,t'}} \cdot \frac{\phi_{rd,tt'}^k}{\lambda_{l,t}^k} \quad (42)$$

$$X_{d,t'}^k = \iota^k \left[ \sum_k \left( w_{d,t'}^{k,G} L_{d,t'}^{k,G} + \sum_l \varphi_{dr,t'}^k \frac{1+\theta}{\theta} w_{r,t'}^{k,G} L_{r,t'}^{k,G} \right) \right] \quad (43)$$

$$\phi_{rd,tt'}^k = \frac{\left[ \sum_{o'} \Omega_{ro',t}(t' - t) \left( w_{r,t'}^{k,G} \tau_{ro',t'}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r,t}^k}{\sum_{r'} \sum_{\tilde{t} \in \mathcal{T}_{-\infty}^{t'}} \left[ \sum_{o'} \Omega_{r'o',\tilde{t}}(t' - \tilde{t}) \left( w_{o',t'}^{k,G} \tau_{o'd,t'}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r',\tilde{t}}^k}, \quad (44)$$

the growth rate of prices is  $g_p = -\frac{1}{\theta} \sum_k \iota^k g^k$  and worker value is  $g_v = \frac{1+\zeta}{\zeta} \frac{1}{\theta} \sum_k \iota^k g^k$ , with  $g^k$  as the Perron-Frobenius root of equation (33), as described in Proposition 2.

The equilibrium conditions on the transition path illustrate the mechanics of innovation in dynamic spatial equilibrium. The evolution of the distribution of workers across regions, sectors, and occupations – given by equations (35) and (36) – along with exogenous research productivities, determine the trajectory of innovation and technology levels given by equations (30) and (37). The trajectory of innovation levels, along with exogenous bilateral diffusion lags, determine the trade equilibrium for each production time, as characterized by equations (38)-(40) and (23). In particular, the market clearing condition pins down contemporaneous production worker wages. The trade equilibrium also yields the probability that goods produced in region  $o$  at time  $t'$  uses ideas discovered in region  $r$  at time  $t$ , given by equation (44). Trajectories of this probability determine incentives to innovate, captured by the value of individual ideas in equation (42) and hence inventor wages in equation (41). In turn, the wages of production workers and inventors determine the incentives of workers to migrate – given by equation (34) – and hence determine the evolution of the distribution of workers across regions, sectors, and occupations.

Notice that the trade equilibrium along the transition path (equations 38-40 and 23) cannot be expressed in terms of time differences. Because ideas have varying applicabilities across regions and over time, trade shares depend on the entire history of past innovations from  $t = -\infty$ , rather than solely on the contemporaneous technology stock. Consequently, two additional sets of variables are required to compute the trade shares and simulate the economy compared to existing quantitative spatial models. First, assuming that the economy is on a balanced growth path for  $t \leq 0$ , the technology stock at  $t = 0$  serves as a sufficient statistic for the trajectory of innovation levels for  $t \leq 0$ . Second, data on wages at  $t = 0$ , along with trade shares and technology levels, allow for the imputation of initial trade costs from equation (79). Consequently, simulating the transition path for  $t \geq 0$  requires only information on changes in trade costs over time like in standard dynamic spatial models, as these changes can be converted to absolute levels given initial trade costs. Note

also that the trade equilibrium does not incorporate detrended variables because production worker wages remain constant on the balanced growth path and can be determined independently of the sectoral price index, which declines over time. This simplification streamlines the numerical algorithm by allowing prices to be detrended *after* computing the trajectory of trade equilibria.

### 4.3 Welfare Impacts of Shocks to Economic Fundamentals

Given the characterizations of the balanced growth and transition paths, I now decompose the welfare impact of the ICT shock – or any arbitrary anticipated sequence of counterfactual changes in fundamentals – into transitory and long-run growth components. Let  $\dot{x}$  denote the counterfactual path and  $\widehat{x}_{t'} = \frac{\dot{x}_{t'}}{\dot{x}_t} = \frac{\dot{x}_{t'}/\dot{x}_t}{x_{t'}/x_t}$  denote counterfactual changes for any variable  $x$ . Define the welfare impact in **market**  $(d, k, h)$  of an anticipated sequence of counterfactual changes in fundamentals from time  $t = 0$  as the compensating variation in consumption for market  $(d, k, h)$ ,  $\log \delta_d^{k,h}$ , given by the following equation:

$$\dot{V}_{d,0}^{k,h} = V_{d,0}^{k,h} + \sum_{t' \in \mathcal{T}_0^\infty} \left( \frac{1}{1+\zeta} \right)^{t'} \log \delta_d^{k,h}.$$

Using these notations and definition of welfare, the impact of an anticipated sequence of counterfactual changes in fundamentals is given by the following corollary:

**Proposition 4.** *Given the transition and balanced growth paths, the **welfare** effects in each market (i.e. region-sector-occupation) of an anticipated counterfactual change in fundamentals is given by:*

$$\log(\delta_d^{k,h}) = \sum_{t' \in \mathcal{T}_{\mathbb{R}^+}} \left( \frac{1}{1+\zeta} \right)^{t'} \log \left( \underbrace{\frac{\widehat{w}_{d,t}^{k,h}}{\widehat{P}_{d,t}}}_{\text{change in future detrended real wages}} \underbrace{\frac{1}{\left( \widehat{\mu}_{dd,t}^{kk} \right)^{1/\Upsilon} \left( \widehat{\mu}_{dd,t}^{kk,hh} | \widehat{\mu}_{dd,t}^{kk} \right)^{v/\Upsilon}}} \right) + \underbrace{\frac{1}{\theta} \sum_k \iota^k (\dot{g}^k - g^k)}_{\text{growth effects}}. \quad (45)$$

while local and aggregate welfare is defined as the population-weighted average of the welfare impacts in the relevant markets.

There are two key differences in this welfare expression compared to Caliendo et al. (2019). First, the option value of moving is determined by both the own-migration share across region-sectors and the conditional own-migration share across occupations, because workers' idiosyncratic preferences for each occupation are correlated across region-sectors. Second, the welfare expression accounts for long run growth, since a counterfactual change in fundamentals affects both transitory and growth outcomes in my model.

## 5 Conclusion

The rise of high-tech clusters over the past half-century in the United States has been a central focus of research on the geography of innovation and a frequent topic in the popular press. Despite extensive attention, the fundamental drivers of this trend have remained elusive, partly due to the wide range of potential explanations.

I first leverage comprehensive data on patents, firms, and inventors from 1976 to 2018 to carefully document when and where innovation became more spatially concentrated. I find that the rising concentration occurred predominantly in high-skill regions and began only after 1990. I then decompose the aggregate trend by technology field and by firm. The results show that the surge in concentration was jointly driven by patent growth in ICT-oriented firms and in high-skill CZ firms. Two distinct aspects of the rise in information and communication technologies (the ICT shock) led to the growth of these firms. First, the sudden arrival of influential ICT breakthroughs from 1985-1993 drove the compositional shift in innovation towards ICT, which is *colocated* with the ICT service sector and concentrated in high-skill CZs. Second, ICT-induced reductions in communication costs enabled high-CZ firms to expand across the country and set up service, logistics and manufacturing plants, increasing their scale and innovation output. Worker sorting toward high-skill CZs in the 1980s facilitated the subsequent growth of these two groups of firms.

To capture the role of each mechanism in shaping the geography of US innovation following the ICT shock, I develop a spatial growth model that integrates endogenous innovation with technology diffusion at the level of *individual ideas*, alongside dynamic worker mobility. The ICT shock enters the model through two exogenous components: a one-time increase in national ICT research productivity in the initial period, and a rise in the national component of bilateral diffusion speeds, reflecting falling communication costs. The model offers analytical characterizations of two key empirical mechanisms: (i) the degree of *colocation* between innovation and production, and; (ii) the *asymmetric scale effect*. These mechanisms determine the wages of innovation workers and, through dynamic mobility, enable *worker sorting* toward high-skill CZs in the 1980s to shape the geography of innovation along the transition path after the ICT shock. A further advantage of the model is its ability to quantify the aggregate consequences of the ICT shock. The steady-state distribution of workers and innovation rates determines the rate at which prices fall and thus the economy's long-run growth rate. Using the model's characterizations of the transition path and balanced growth path, I analytically decompose the welfare impact of the ICT shock into its transitory and long-run growth components.

More broadly, my model leverages the probabilistic microfoundations of the Eaton-Kortum structure (Kortum, 1997; Eaton and Kortum, 1999, 2001, 2002; Bernard et al., 2003; Eaton and Kortum, 2024) to introduce *endogenous and directed innovation* into modern dynamic spatial models in a highly tractable manner, offering methodological tools to examine the causes and consequences of the geography of innovation in different contexts. This paper thus establishes the foundation for a broader research agenda on the spatial and network aspects of innovation.

## References

- Acemoglu, D. (1998). Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *The Quarterly Journal of Economics*, 113(4):1055–1089.
- Acemoglu, D. (2002). Directed Technical Change. *Review of Economic Studies*, 69(4):781–809.
- Acemoglu, D. (2007). Equilibrium Bias of Technology. *Econometrica*, 75(5):1371–1409.
- Allen, T. and Arkolakis, C. (2014). Trade and the Topography of the Spatial Economy. *Quarterly Journal of Economics*, (2002):1085–1139.
- Alvarez, F. and Lucas, R. E. (2007). General equilibrium analysis of the Eaton-Kortum model of international trade. *Journal of Monetary Economics*, 54(6):1726–1768.
- Andrews, M. J. and Whalley, A. (2022). 150 Years of the Geography of Innovation. *Regional Science and Urban Economics*, 94.
- Arcidiacono, P., Bayer, P., Blevins, J. R., and Ellickson, P. B. (2016). Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition. *The Review of Economic Studies*, 83(3):889–931.
- Autor, D. H. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5):1553–97.
- Ben-Akiva, M. and Francois, B. (1983). Mu-homogenous generalized extreme value model. Technical report, Working paper, Department of Civil Engineering, MIT.
- Bernard, A. B., Eaton, J., Jensen, J. B., and Kortum, S. (2003). Plants and Productivity in International Trade. *American Economic Review*, 93(4):1268–1290.
- Bikard, M. and Marx, M. (2020). Bridging Academia and Industry: How Geographic Hubs Connect University Science and Corporate Technology. *Management Science*, 66(8):3425–3443.
- Bresnahan, T. and Greenstein, S. (1996). Technical Progress and Co-Invention in Computing and in the Uses of Computers. *Brookings Papers on Economic Activity. Microeconomics*, 1996:1–83.
- Buera, F. J. and Oberfield, E. (2020). The Global Diffusion of Ideas. *Econometrica*, 88(1):83–114.
- Cai, J., Li, N., and Santacreu, A. M. (2022). Knowledge Diffusion, Trade, and Innovation across Countries and Sectors. *American Economic Journal: Macroeconomics*, (1):104–145.
- Cai, S., Caliendo, L., Parro, F., and Xiang, W. (2025). Mechanics of Spatial Growth. *Working Paper*.
- Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock. *Econometrica*, 87(3):741–835.

- Caliendo, L. and Parro, F. (2015). Estimates of the Trade and Welfare Effects of NAFTA. *Review of Economic Studies*, 82(1):1–44.
- Choi, K.-H. and Moon, C.-G. (1997). Generalized extreme value model and additively separable generator function. *Journal of Econometrics*, 76(1-2):129–140.
- Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2018). The Geography of Development. *Journal of Political Economy*, 126(3):903–983.
- Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000. *American Economic Review*, 106(3):479–524.
- Diamond, R. and Gaubert, C. (2022). Spatial Sorting and Inequality. *Annual Review of Economics*.
- Dreisigmeyer, P., Goldschlag, N., Krylova, M., Ouyang, W., and Perlman, E. (2018). Building a Better Bridge: Improving Patent Assignee-Firm Links. *CES Technical Notes Series*, (1).
- Eaton, J. and Kortum, S. (1999). International technology diffusion: Theory and measurement. *International Economic Review*, 40(3):537–570.
- Eaton, J. and Kortum, S. (2001). Technology, trade, and growth: A unified framework. *European Economic Review*, 45:742–755.
- Eaton, J. and Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779.
- Eaton, J. and Kortum, S. (2024). Technology and the Global Economy. *Annual Review of Economics*, 16(Volume 16, 2024):79–104.
- Eckert, F., Fort, T., Schott, P., and Yang, N. (2020). Imputing Missing Values in the US Census Bureau's County Business Patterns. Technical Report w26632, National Bureau of Economic Research, Cambridge, MA.
- Ellison, G. and Glaeser, E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5):889–927.
- Feldman, M. P. and Kogler, D. F. (2010). *Stylized Facts in the Geography of Innovation*, volume 1. Elsevier B.V.
- Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, J. R., Warren, D., Breton, E., Cooper, G., Drew, J. A. R., Richards, S., Van Riper, D., and Williams, K. C. (2025). IPUMS, Current Population Survey: Version 13.0.
- Fort, T. C., Keller, W., Schott, P. K., Yeaple, S., and Zolas, N. (2020). Colocation of Production and Innovation: Evidence from the United States. *Working Paper*.
- Foster, L., Haltiwanger, J., and Krizan, C. (2001). Aggregate Productivity Growth: Lessons from Microe-

- conomic Evidence. In *New Developments in Productivity Analysis*, page w6803. National Bureau of Economic Research, Cambridge, MA.
- Greenstein (2015). *How the Internet Became Commercial: Innovation, Privatization, and the Birth of a New Network*. Princeton University Press.
- Greenstein, S. and Prince, J. (2006). The Diffusion of the Internet and the Geography of the Digital Divide in the United States. Technical Report w12182, National Bureau of Economic Research, Cambridge, MA.
- Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. *NBER Working Paper*.
- Hemous, D., Lepot, S., Sampson, T., and Schärer, J. (2025). Trade, Innovation and Optimal Patent Protection. *Working Paper [Revise and Resubmit, American Economic Review]*.
- Hsieh, C.-T. and Rossi-Hansberg, E. (2023). The Industrial Revolution in Services. *Journal of Political Economy Macroeconomics*, 1(1):3–42.
- Jiang, X. (2024). Information and Communication Technology and Firm Geographic Expansion. *Working Paper [Revise and Resubmit, Journal of Political Economy: Macroeconomics]*.
- Jones, C. I. (1995). R&D-Based Models of Economic Growth. *Journal of Political Economy*, 103(4):759–784.
- Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2021). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights*, 3(3):303–20.
- Kerr, W. R. and Fu, S. (2008). The survey of industrial R&D—patent database link project. *The Journal of Technology Transfer*, 33(2):173–186.
- Kleinman, B. (2025). Wage Inequality and the Spatial Expansion of Firms. *Working Paper [Revise and Resubmit, Quarterly Journal of Economics]*.
- Kleinman, B., Liu, E., and Redding, S. J. (2023). Dynamic Spatial General Equilibrium. *Econometrica*, 91(2):385–424.
- Kortum, S. (1997). Research, Patenting, and Technological Change. *Econometrica*, 65(6):1389–1419.
- Krugman, P. R. (1992). *Geography and Trade*. MIT Press.
- Li, G. C., Lai, R., D'Amour, A., Doolin, D. M., Sun, Y., Torvik, V. I., Yu, A. Z., and Lee, F. (2014). Disambiguation and co-authorship networks of the U.S. patent inventor database (1975-2010). *Research Policy*, 43(6):941–955.
- Lind, N. and Ramondo, N. (2023). Global Innovation and Knowledge Diffusion. *American Economic Review: Insights*, 5(4):494–510.

- Lind, N. and Ramondo, N. (2024). Global knowledge and trade flows: Theory and measurement. *Journal of International Economics*, 151:103960.
- Liu, J. (2024). Multinational Production and Innovation in Tandem. *Working Paper*.
- McFadden, D. (1978). *Modeling the Choice of Residential Location*, pages 75–96. North Holland, Amsterdam.
- Moretti, E. (2013). Real Wage Inequality. *American Economic Journal: Applied Economics*, 5(1):65–103.
- Moretti, E. (2021). The Effect of High-Tech Clusters on the Productivity of Top Inventors. *American Economic Review*, 111(10):3328–3375.
- Mowery, D. C. and Simcoe, T. (2002). Is the Internet a US invention?—an economic and technological history of computer networking. *Research Policy*, 31(8):1369–1387.
- Ramondo, N. and Rodríguez-Clare, A. (2013). Trade, Multinational Production, and the Gains from Openness. *Journal of Political Economy*, 121(2):273–322.
- Schmoch, U. (2008). Concept of a Technology Classification for Country Comparison. *Final report to the World Intellectual Property Organization*, (June):1–15.
- Schroeder, J., Van Riper, D., Manson, S., Knowles, K., Kugler, T., Roberts, F., and Ruggles, S. (2025). National Historical Geographic Information System: Version 20.0.
- Somale, M. (2021). Comparative Advantage in Innovation and Production. *American Economic Journal: Macroeconomics*.
- Xiang, W. (2023). Clean Growth and Environmental Policies in the Global Economy. *Working Paper*.

## A Data Details for the Empirical Analysis

Here I provide detailed information about the data sources I used to measure and document fundamental trends in the geography of innovation in the United States, emphasizing my newly constructed six broad fields from technology classes under the Cooperative Patent Classification (CPC) scheme.

### A.1 Patent Data and Cleaning Procedures

Patent data comes from bulk files from the US Patent Trademark Office (USPTO) and US PatentsView (USPV) (<https://patentsview.org/>). The datasets contain the universe of patents granted in the US (by the USPTO) from 1976-2022. I use only utility patents – which comprise about 95% of all granted patents from 1976 – as they are patents for inventions and provide the most appropriate measure of technological advancements, as opposed to other types of patents such as design, plant, and defensive patents. As explained by the USPTO Patent Technology Monitoring Team (PTMT)<sup>27</sup>:

*Most analyses of technological activity that incorporate patent data will focus on the activity of utility patents, also known as “patents for inventions”. Since design patents are granted for ornamental designs for articles of manufacture and not for inventions, they are usually perceived to have a lesser relationship with technological activity. Similarly, statutory invention registrations and defensive publications do not convey patent protection to disclosed inventions and may have a lesser relationship with technological activity. Plant patents may or may not disclose an invention resulting from technological activity; however, plant patents are numerically small relative to utility patents and are usually handled and analyzed separately.*

Inventors, assignees, and locations are disambiguated by USPV – meaning each inventor, assignee, and city-state pair has a unique ID over time – using the latest machine learning techniques, improving on the algorithms used in Li et al. (2014) and to produce the Connecting Outcome Measures in Entrepreneurship, Technology, and Science (COMETS) database. Cities and states of all inventors living in the US are provided. I use the Google Maps API to geocode all locations and assign them to counties, and standard publicly-available crosswalks to convert counties to commuting zones (CZs) [using Autor and Dorn (2013)], core-based statistical areas (CBSAs), and combined statistical areas (CSAs). The high spatial resolution of inventor locations allows me to conduct my analysis at different geographical scales.

I define the patent year as the application year since that is the closest to when the invention was produced, as there are often lags of several years between when a patent was applied and when it was granted. I assign patent shares, citations made, and citations received equally across all coinventors on a patent. I keep only patents where at least one inventor lives in the US.

---

<sup>27</sup>The USPTO PTMT explanation on the conventions of treating different types of patents can be found at <https://www.uspto.gov/web/offices/ac/ido/oeip/taf/reports.htm>. Should this or any of the subsequent web links become inactive, PDF copies of the contents of the archived websites that includes the date of access will be provided upon request.

## A.2 Classification of Patents into Technology Fields and Subfields

Patents are assigned to 3-digit technology classes, 4-digit subclasses, and further divided into groups within subclasses under the International Patent Classification (IPC) and Cooperative Patent Classification (CPC) systems. The IPC provides a hierarchical system of language independent symbols for the classification of patents into eight sections with approximately 70,000 subdivisions. The IPC was developed under the 1971 Strasbourg Agreement and provides a common classification for patents filed in different patent offices around the world. The CPC is a unified system developed jointly by the United States Patent Trademark Office (USPTO) and European Patent Office (EPO) on 2010 to provide a common, internationally compatible classification system that provides more groups and subgroups relative to the IPC. For patents granted in the US, the CPC supercedes the US Patent Classification (USPC) developed by the USPTO and the NBER patent classification developed by Hall et al. (2001). The USPTO no longer provides the USPC and NBER patent classifications for patents granted after 2015.

Most patents have more than one field classification to facilitate easier searches to prior art. Nonetheless, under the CPC, each patent has a unique primary field classification, denoted as the “first” position. Other classifications are denoted as having a “later” position. USPTO Guideline 905.03(a)III.A.(C)<sup>28</sup> states:

*There is one and only one “first” position attribute per patent family. The first attribute is associated with the invention symbol that most completely covers the technical subject matter of the disclosed invention. The first position symbol is identified as the first mandatory symbol listed on the classification form.*

The USPTO provides a bulk file of all granted patents from 1790 with their current CPC classes and position of each CPC class<sup>29</sup>. I use the 08/02/2022 version, downloaded on 09/09/2022<sup>30</sup>. I use the primary CPC class of each patent. In rare cases where there are multiple CPC classes listed as the “first” position, I use the class that is listed first. In rare cases where the CPC class listed in the “first” position is only meant as an additional classification as noted under the details of the CPC class in their classification documentation<sup>31</sup>, I use the class in the “later” position that is listed first.

Under the IPC and CPC, the technology subclasses are grouped into 8 broad sections. These broader sections facilitate the allocation of patents to different examiner units at patent offices. **However, some sections contain patents from highly disparate economic fields. For example, Section A includes patents for medical technologies and amusement parks.** To provide more consistent categories with similar sizes, the World Intellectual Property Organization (WIPO) developed a mapping from the 4-digit technology subclasses and groups in the IPC/CPC to 35 fields (Schmoch, 2008). This WIPO report also provides 6

<sup>28</sup>The USPTO guidelines on the CPC scheme can be found at <https://www.uspto.gov/web/offices/pac/mpep/s905.html>

<sup>29</sup>The bulk files containing the CPC classification of every patent can be found at <https://bulkdata.uspto.gov/data/patent/classification/cpc/>

<sup>30</sup>Although USPV provides CPC classes and technology fields for most patents granted since 1976, data checks reveal that the 2022 vintages of the USPV datasets contain errors in the CPC classes and WIPO technology fields for approximately 3–5% of patents.

<sup>31</sup>The CPC classification documentation can be found at <https://www.uspto.gov/web/patents/classification/cpc/html/cpc.html>.

broad categories, but the broad category “Instruments” is difficult to interpret and the “Electrical Engineering” category includes fields in information technology as well as older electrical machinery. Thus, I provide an alternative mapping of these 35 WIPO fields into 6 broad categories: (1) Physics, Electrical Engineering & Electronics; (2) Information Technology; (3) Chemistry; (4) Biology & Medicine; (5) Mechanical Engineering; (6) Civil Engineering & Consumer Goods. These categories are similar in scope to the older field classifications in the NBER and COMETS datasets and may be seen as an updated version of them.

The IPC/CPC class to WIPO field concordance is primarily at the level of 4-digit technology subclasses, but there are 6 subclasses with different groups that map to different WIPO fields – A61K, B01D, C13B, E01F, G01N, H04N – due to significantly different subject matter across groups within these subclasses. For example, within the subclass A61K, there are patents for cosmetics and patents for medical technologies. Thus, I provide a further decomposition for these 6 subclasses that corresponds to the WIPO field they are assigned to. My added subclasses are: A61K-14; A61K-16; B01D-23; B01D-24; C13B-18; C13B-29; E01F-24; E01F-35; G01N-10; G01N-11; H04N-2; H04N-3; H04N-4. The numbers after the hyphen refer to the WIPO fields that groups within each subclass are assigned to. Thus, my data has 586 technology classes, comprising the 579 CPC/IPC subclasses along with this decomposition of 6 subclasses. In this paper, I refer to the 35 World Intellectual Property Organization (WIPO) fields as “subfields”, my 6 broad categories as “fields”, and the 579 CPC/IPC subclasses and 7 additional group categories as “classes”.

### A.3 Data Sources for Regional Outcomes

I use data on county-level educational attainment for each decade from the Economic Research Service (ERS) of the US Department of Agriculture (USDA) to construct the college ratio and percent of workers that are college educated for each CZ in 1970, 1980, 1990, 2000, 2010 and 2017. The 1970, 1980, 1990 and 2000 data are constructed from the Decennial Censuses of Population and equivalent to the 5% samples from the IPUMS used by Moretti (2013); Diamond (2016) to construct college percent and ratio respectively for metropolitan statistical areas (MSAs) in 1980 and 2000. The 2010 and 2017 data are constructed from 5-year averages from the annual American Community Survey (ACS). Data from the IPUMS NHGIS (Schroeder et al 2025) are identical.

I obtain annual data on the college-educated and non-college educated workers in each IND1990 industry from IPUMS CPS. I define the ICT sector to include

I obtain annual county-level employment by NAICS industry for 1975-2018 from the harmonized County Business Patterns (CBP) Database produced by Eckert et al. (2020). They develop a linear programming method to impute suppressed county-industry-year cells in raw CBP files released by the US Census Bureau and show that total non-agricultural employment from their dataset is highly correlated with the panel from the Longitudinal Business Database (LBD). I use their county-industry employment panel to construct my primary measures for the annual industrial composition of each CZ and the United States, such as the employment share in the ICT service sector. Building on Fort et al. (2020), I define the ICT service sector to include the following industries in the Information Sector (NAICS 51): Software Publishers (5112);

Telecommunications (517); Data Processing, Hosting, and Related Services (518).

I also obtain annual county-level data on income and resident population for 1969-2018 from the Bureau of Economic Analysis to construct basic annual measures of each CZ such as population, population density, and income per capita.

## B Additional Empirical Results

Here I present supplementary empirical results on the main facts and drivers of the post-1990 rise in the spatial concentration of US innovation. Appendix B.1 presents robustness checks on Fact 1, Appendix B.2 presents additional results and robustness checks on the field-level decomposition (Facts 2 and 3), Appendix B.3 presents firm-level decompositions for each component from the field-level decomposition (Facts 4 and 5), and Appendix B.4 presents details of the NSFNET.

### B.1 Fact 1: Rising Concentration of Innovation in High-Skill CZs from 1990

Figure 11 presents robustness checks on the trend in the aggregate spatial concentration of innovation in Figure 1 in the main paper. Innovation became more concentrated in space only after 1990, whether I drop top patenting CZs (left graph) or use alternative measures of spatial concentration such as the Herfindahl index, coefficient of variation, simplified Ellison-Glaeser measure of patent shares (middle graph), or the annual share of patents produced by the top 10 CZs (right graph).

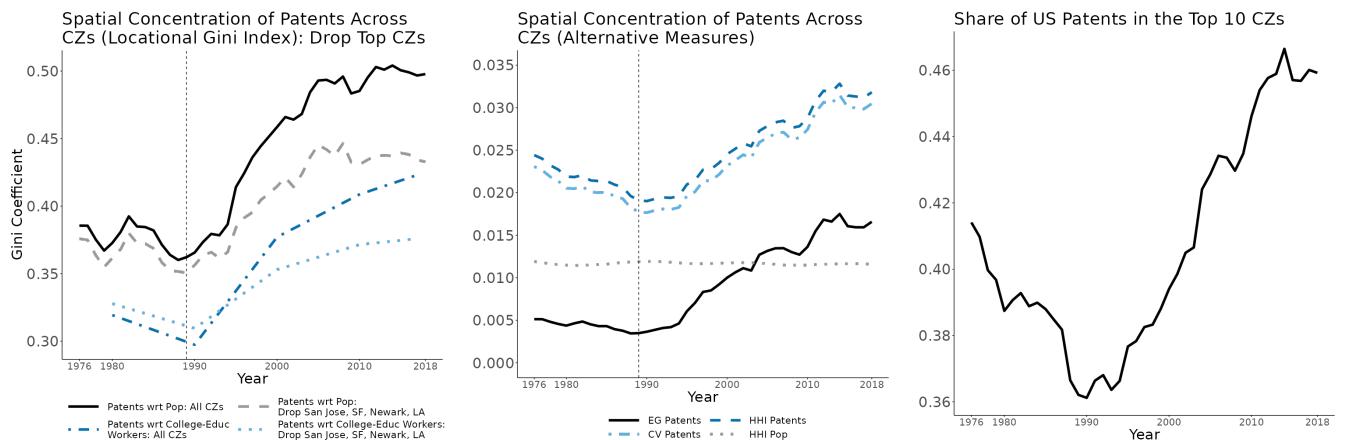


Figure 11: Robustness of trends in the spatial concentration of patents across CZs. The top left graph plots trends in the locational Gini index of patents with respect to population and college-educated workers after removing top CZs in patents: San Jose, San Francisco, Newark, and Los Angeles. The middle graph plots trends in alternative measures of spatial concentration, such as the Herfindahl index (blue dashed line), coefficient of variation (light blue dotdashed line), and simplified Ellison-Glaeser measure (black solid line) of patent shares. The right graph plots trends in the annual share of patents produced by the top 10 CZs.

Figure 12 depicts the geography of changes in the patent-to-population share ratio between 5-year averages around 1990 and 2015. Figure 13 shows how the geography of these changes relate to the rise in the Gini

coefficient since 1990. The left panel displays the patent-to-population share ratio across regions in 1990, while the right panel shows that regions with a higher ratio in 1990 generally maintained an even higher ratio in 2015. Conversely, regions with a lower ratio in 1990 experienced an even lower ratio by 2015.

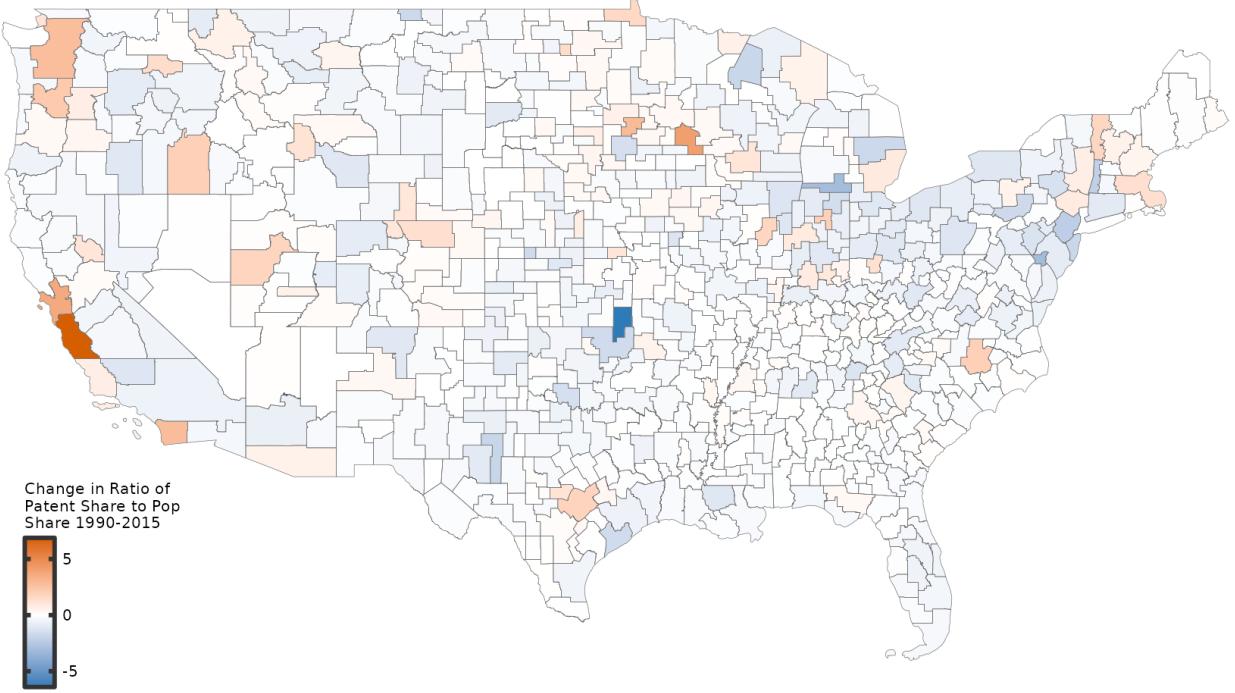


Figure 12: Changes in the geography of innovation intensity from 1990 to 2015, measured using changes in the patent-to-population share ratio on a pseudo-log scale.

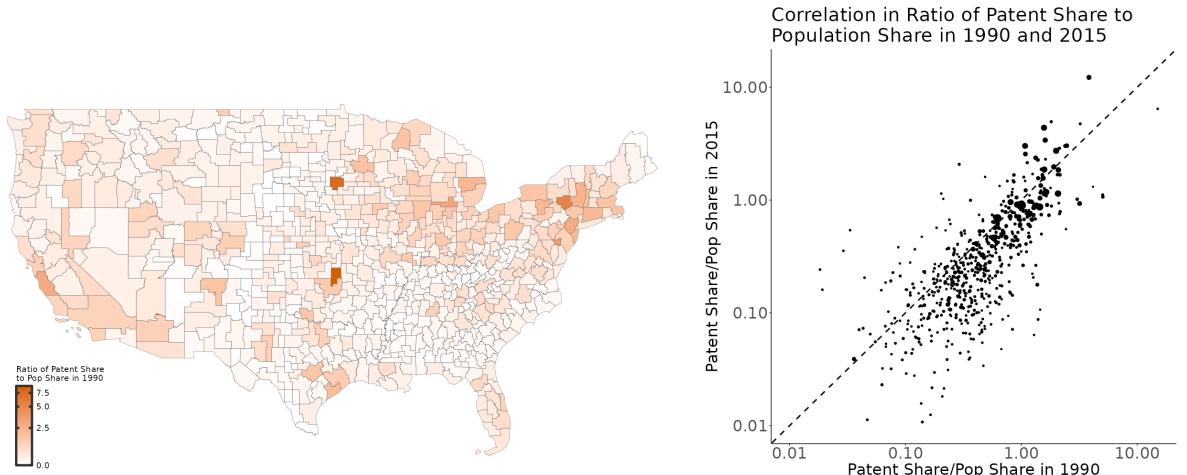


Figure 13: The geography of the patent-to-population share ratio in 1990 (left) and its correlation with the same ratio in 2015 (right)

Table 4 lists the top 15 CZs by changes in patent-to-population share between 5-year averages around 1990 and 2015. Apart from well-known superstar cities such as San Jose, San Francisco, San Diego, Seattle and Boston, CZs like Portland, Boise, Wayne, Provo and Fort Collins also became some of the most innovative

regions in the US from 1990-2015. Despite narratives of the apparent collapse of innovation in Rochester in several papers and the popular press, the region actually experienced the second greatest increase in patenting intensity from 1990-2015.

	CZ	Change in patent-to-population share 1990-2015	Patent-to-population share	Rank of patent-to-population share	Patent Share	Patents	Population
1	San Jose	9.494	12.421	1	0.104	14581.5	2684061
2	Rochester	3.608	4.618	2	0.004	531.4	263093
3	San Francisco	3.029	4.440	3	0.073	10153.8	5229363
4	San Diego	2.094	3.430	5	0.035	4882.0	3254818
5	Seattle	2.063	2.955	8	0.043	5974.0	4622131
6	Portland	1.618	2.523	11	0.018	2521.7	2285666
7	Boise	1.411	1.962	18	0.004	604.7	704711
8	Raleigh	1.398	2.282	13	0.015	2064.7	2068936
9	Wayne	1.396	2.303	12	0.001	132.5	131541
10	Austin	1.258	2.959	7	0.019	2660.2	2055406
11	Provo	1.232	1.763	19	0.003	467.1	605648
12	Bloomington	1.186	1.587	25	0.001	157.2	226557
13	Burlington	1.167	2.224	14	0.002	333.9	343199
14	Boston	0.885	2.730	9	0.046	6488.4	5434761
15	Fort Collins	0.871	1.756	20	0.004	495.9	645656
	Oneonta	-0.944	4.414	4	0.002	300.3	155554
	Albany	0.466	2.959	6	0.010	1461.1	1128888
	Elmira	0.341	2.627	10	0.003	401.9	349814
	Poughkeepsie	0.554	2.185	15	0.006	897.9	939410
	Minneapolis	0.192	2.021	16	0.021	2989.4	3381353
	Detroit	0.588	1.967	17	0.032	4510.1	5243719

Table 4: The top 15 CZs by change in ratio of patent-to-population share from 1990-2015 with a population of at least 100,000 in 2015. To minimize the effect of idiosyncratic fluctuations, I take five-year averages around 1990 and 2015. I append this list with the 6 CZs that are in the top 20 by ratio of patent-to-population share in 2015 but did not experience the greatest increase from 1990 to 2015.

Figure 14 plots trends in the annual elasticity of CZ patents per capita with respect to the 1990 population (left graph) and population density (right graph). The left graph shows that while the annual elasticity of patents per capita with respect to the 1990 population increased moderately from 1990-2018, these increases did not overcome the decrease in annual elasticity from 1976-1990. The right graph shows that annual elasticity of patents per capita with respect to the 1990 population density did not increase after 1990.

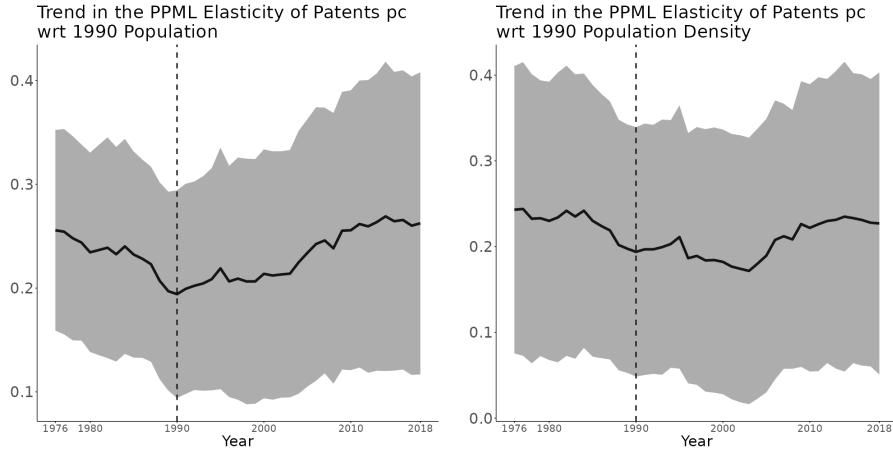


Figure 14: Absence of evidence on patent sorting to large or dense CZs after 1990. The left panel displays the trend in the annual elasticity of CZ patents per capita with respect to the 1990 population while the right panel displays an analogous trend with respect to the 1990 population density. The confidence intervals reflect the statistical uncertainty of the slope estimates for each year given heteroskedastic-robust standard errors.

## B.2 Field Decomposition of the Rising Spatial Concentration of Innovation

Here I present additional results on the field-level decomposition (**Facts 2 and 3** in the main text). Appendix B.2.1 presents trends in the ICT and non-ICT patent elasticity (i.e. the elasticity of ICT versus non-ICT patents per capita with respect to the 1990 college ratio). Appendix B.2.2 provides a formal decomposition of the post-1990 rise in patent elasticity into within-field and cross-field components. Appendix B.2.3 presents an analogous decomposition of the post-1990 rise in the locational Gini of patents per capita. Appendix B.2.4 examines within-field changes in both the locational Gini and patent elasticity after 1990 in greater detail. Appendix B.2.5 examines the robustness of the field-level trends and decomposition to dropping top patenting CZs. Appendix B.2.6 presents analogous trends and decompositions across subfields.

### B.2.1 Trends on ICT versus non-ICT Patents

In the main text, the right panel of Figure 3 presents trends in patent elasticity by technology field. To better illustrate how ICT differs from other fields, I now aggregate all non-ICT fields into a single category. The left panel of Figure 15 plots the 1990 patent elasticity for ICT vs non-ICT, i.e. the correlation between patents per capita and the college ratio in 1990, where each dot represents a CZ and the size of the dot reflects the CZ's population. The graph shows that in high-skill CZs, ICT and non-ICT patents per capita are similar, whereas in lower-skill CZs, non-ICT patents far exceeds that of ICT. This highlights the stronger concentration of ICT innovation in high-skill CZs.

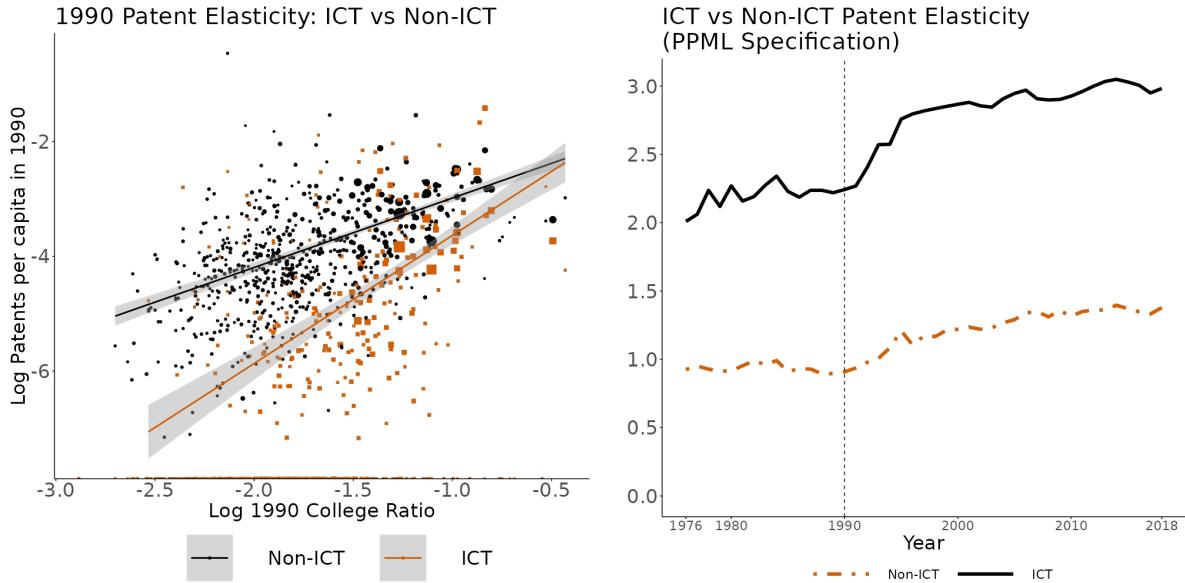


Figure 15: Trends in the concentration of ICT versus non-ICT patents in high-skill CZs. The left panel displays the correlation between patents per capita and the college ratio in 1990, where each dot represents a CZ and the size of the dot reflects the CZ's population. The gradients represent the ICT and non-ICT patent elasticity respectively in 1990. The right panel shows trends in the patent elasticity for ICT versus non-ICT.

### B.2.2 Decomposition of the Rising Patent Elasticity $\alpha_t$ after 1990 into Within-Field and Cross-Field Components

More formally, I decompose the post-1990 increase in the aggregate annual elasticity,  $\alpha_t$ , of CZ patents per capita with respect to the 1990 college ratio into within-field and cross-field components as follows:

$$\alpha_{t^*} - \alpha_{1990} = \sum_{t=1991}^{t=t^*} \Delta \alpha_t = \sum_{t=1991}^{t=t^*} \left[ \underbrace{\sum_k \bar{\alpha}_{k,t} \Delta s_{k,t}}_{\text{changes in field composition}} + \underbrace{\sum_k \bar{s}_{k,t} \Delta \alpha_{k,t}}_{\text{within-field changes}} + \underbrace{\Delta \left( \alpha_t - \sum_k s_{k,t} \alpha_{k,t} \right)}_{\text{residual: changes in the colocation of fields}} \right] \quad (46)$$

where  $\alpha_{k,t}$  is the annual elasticity of CZ patents per capita in field  $k$  and year  $t$  with respect to the 1990 college ratio, and  $\bar{x}_t = \frac{x_t + x_{t-1}}{2}$  and  $\Delta x_t = x_t - x_{t-1}$  denote the average and change of any variable  $x$  between  $t-1$  and  $t$ . The first term reflects the impact of changes in the field composition of US patents. This term is positive if patents are increasingly produced in fields that are more spatially concentrated in high-skill cities. The second term captures the role of changes in the spatial concentration of patents in high-skill cities within fields. The third term accounts for changes in the colocation of fields in high-skill cities, measured by differences in the overall elasticity of patents per capita against the 1990 college ratio from a weighted mean of the field-specific elasticities.

The left panel of Figure 16 plots trends in this decomposition, while the middle and right panels further decompose the cross-field and within-field components respectively into contributions from individual fields. The graphs in the left and middle panels show that 59% of the increase in the overall elasticity is attributed to cross-field changes: 53% of the increase in the overall elasticity is explained by the rising share of ICT patents and 6% from the rising share of Biology patents. The right panel shows that 14% of the increase in the overall elasticity is explained by increases within ICT and 27% by increases across each of the non-ICT fields.

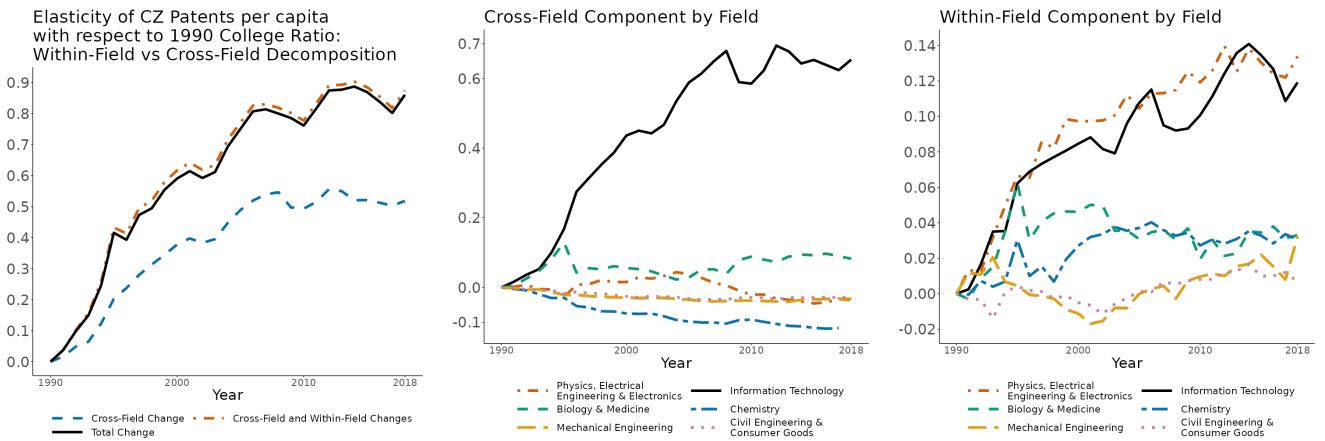


Figure 16: Trends in the decomposition of the aggregate elasticity of CZ patents per capita with respect to the 1990 college ratio. The left panel decomposes the aggregate elasticity into within versus cross field components, and the middle and right panels further break down the cross-field and within-field components respectively into contributions from individual fields.

### B.2.3 Trends on the Rising Locational Gini after 1990 and its Decomposition into Within-Field and Cross-Field Components

A more complete understanding of the trends in the spatial concentration of patents across fields requires examining how the Gini coefficient of patents per capita has evolved over time for each field. For example, patents in some fields may be as concentrated as ICT patents but in different sets of CZs. The left panel of Figure 17 shows trends in the Gini coefficient of patents per capita by technology field, while the right panel displays the same measure for ICT versus non-ICT patents. These trends confirm that ICT patents have consistently been more spatially concentrated than non-ICT patents, and that spatial concentration increased for both groups after 1990.

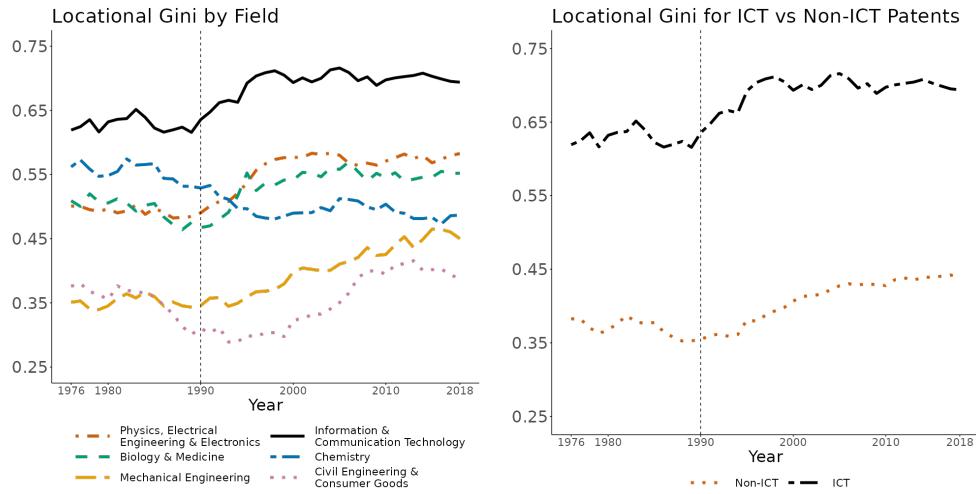


Figure 17: Trends in the Gini coefficient of patents per capita by field (left) and by ICT versus non-ICT (right).

More formally, I decompose changes over time in the overall Gini coefficient  $G$  of patents per capita from 1990-2018 into within-field, cross-field, and field colocation components:

$$G_{t^*} - G_{1990} = \sum_{t=1991}^{t=t^*} \Delta G_t = \sum_{t=1991}^{t=t^*} \left[ \underbrace{\sum_f \bar{G}_{f,t} \Delta s_{f,t}}_{\text{changes in field composition}} + \underbrace{\sum_f \bar{s}_{f,t} \Delta G_{f,t}}_{\text{within-field changes}} + \underbrace{\Delta \left( G_t - \sum_f s_{f,t} G_{f,t} \right)}_{\text{changes in the colocation of fields}} \right]$$

where  $G_{f,t}$  is the Gini coefficient of patents in field  $f$  with respect to population across CZs in year  $t$ ,  $s_{f,t}$  is the share of US patents in field  $f$  in year  $t$ , and  $\bar{x}_t = \frac{x_t + x_{t-1}}{2}$ ,  $\Delta x_t = x_t - x_{t-1}$  for any variable  $x$ . The first term captures the role of changes in the field composition of US patents. This term is positive if patents are increasingly produced in fields that are more spatially concentrated. The second term captures the role of changes in the spatial concentration of patents within fields. The third term captures the role of changes in the colocation of fields, measured by differences in the overall locational Gini coefficient of all CZ patents with respect to CZ population from a weighted mean of the field-specific locational Gini coefficients.

The left panel of Figure 18 plots trends in the decomposition of the rise in the spatial concentration of

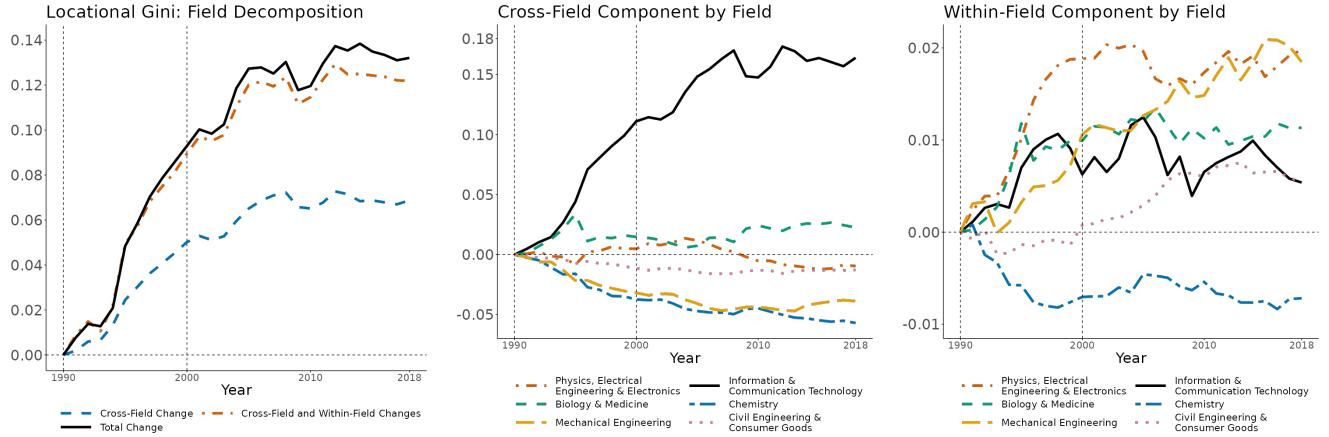


Figure 18: Trends in the decomposition of changes in the locational Gini coefficient of patents with respect to population relative to 1990. The left graph plots trends in the overall decomposition into within-field, across-field, and field colocation components while the middle and right graphs further decompose the cross-field and within-field components respectively into the contribution of each individual field.

innovation from 1990-2018. Over this period, 52% of the increase in the Gini coefficient of patents per capita is explained by changes in the field composition of patents (the cross-field component), 40% by rising spatial concentration within fields (the within-field component), and 8% by the increasing colocation of patents across different fields.<sup>32</sup> The middle panel further decomposes the cross-field component into the contribution of each individual field and shows that nearly the entire effect is driven by the rising share of ICT patents. In particular, 4% of the overall rise is due to increased concentration within ICT, while 36% can be attributed to increased concentration within the non-ICT fields of *Biology & Medicine*; *Physics, Electrical Engineering & Electronics*; *Mechanical Engineering*; and *Civil Engineering & Consumer Goods*.

#### B.2.4 Trends Within Fields

The above decompositions show that the compositional shift in patents toward ICT accounts for 46% of the post-1990 rise in the overall Gini coefficient of patents per capita and 53% of the corresponding rise in patent elasticity. To examine the remaining increase – driven by within-field changes – more closely, Figure 19 replicates the trends in the Gini coefficient by field (from the left panel of Figure 17), and the patent elasticity by field (from the right panel of Figure 3 in the main text), normalizing levels in 1990 to facilitate comparison across fields in the changes over time.

The left panel of Figure 19 shows that the Gini coefficient of patents per capita increased between 1990 and 2018 across all fields except *Chemistry*. Notably, the spatial concentration of patents in *ICT*; *Physics, Electrical Engineering & Electronics*; and *Biology & Medicine* rose sharply during the 1990-2000 period. As shown in the right panel, these fields also exhibited significant increases in patent elasticity over the same period, indicating a growing concentration specifically in high-skill CZs. By contrast, *Mechanical Engineering* patents became more concentrated in manufacturing hubs during the 1990s, with a shift toward

<sup>32</sup>Most of the rise between 1990 and 1995 is accounted for by changes in field composition, whereas the subsequent increase is more evenly split between the cross-field and within-field components.

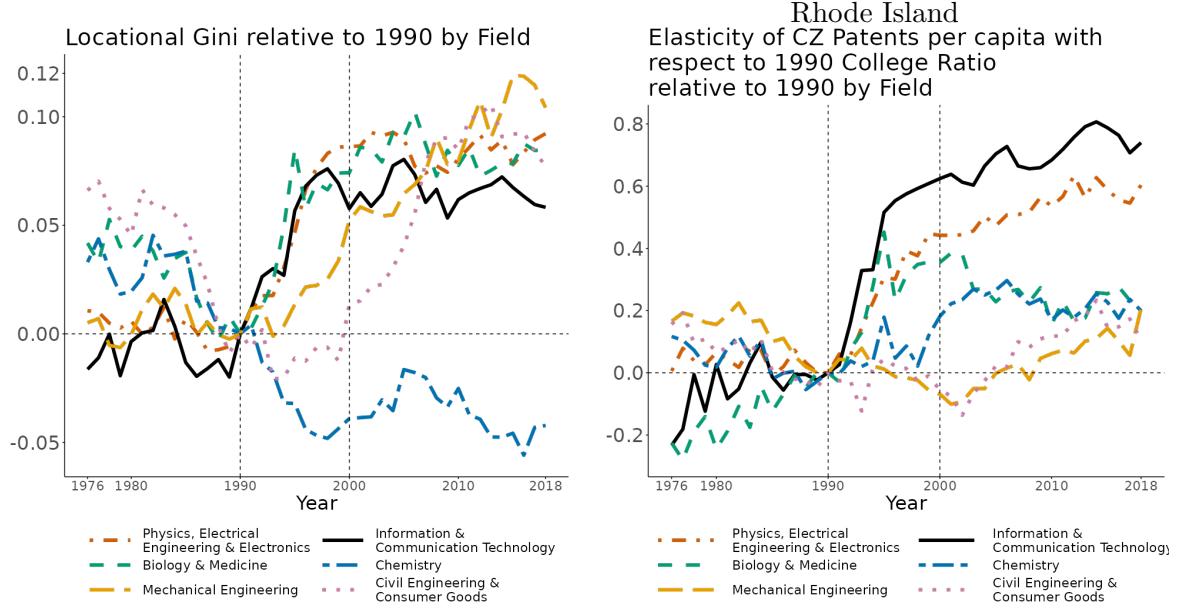


Figure 19: Trends in changes relative to 1990 of the locational Gini coefficient (left) and elasticity with respect to 1990 college ratio (right) of patents per capita by technology field.

high-skill CZs emerging only in the early 2000s. The yellow long-dashed line in the left panel shows that the Gini coefficient for *Mechanical Engineering* rose gradually from 1990 to 2018, while the right panel indicates that its patent elasticity increased only after 2002.<sup>33</sup> Finally, as indicated by the pink dotted line in both panels, the spatial concentration of *Civil Engineering & Consumer Goods* patents remained relatively stable during the 1990s, but increased substantially in high-skill CZs starting in the early 2000s.

### B.2.5 Trends and Decomposition by Field without Top Patenting CZs

One potential concern is that the observed trends in the spatial concentration of innovation are driven primarily by Silicon Valley (San Jose and San Francisco). While I find that Silicon Valley accounts for much of the rise in the locational Gini coefficient after the late 1990s, it does not significantly influence Facts 2 through 4 in the main text.

Figure 20 plots trends in the share of annual patents, the locational Gini coefficient of patents per capita, and patent elasticity by technology field, excluding Silicon Valley. The left panel shows that the compositional shift toward ICT remains largely unchanged. The middle panel confirms that ICT patents continue to exhibit higher spatial concentration than other fields and became more concentrated between 1990 and 2000. However, excluding Silicon Valley leads to a modest decline in ICT spatial concentration after 2000. The right panel shows that trends in patent elasticity are largely unaffected by the exclusion of Silicon Valley, suggesting that the increasing concentration of innovation in high-skill CZs reflects broader national

<sup>33</sup>Supporting evidence for the rising concentration of Mechanical Engineering patents in manufacturing hubs during the earlier period is available upon request.

patterns rather than being driven solely by this region.

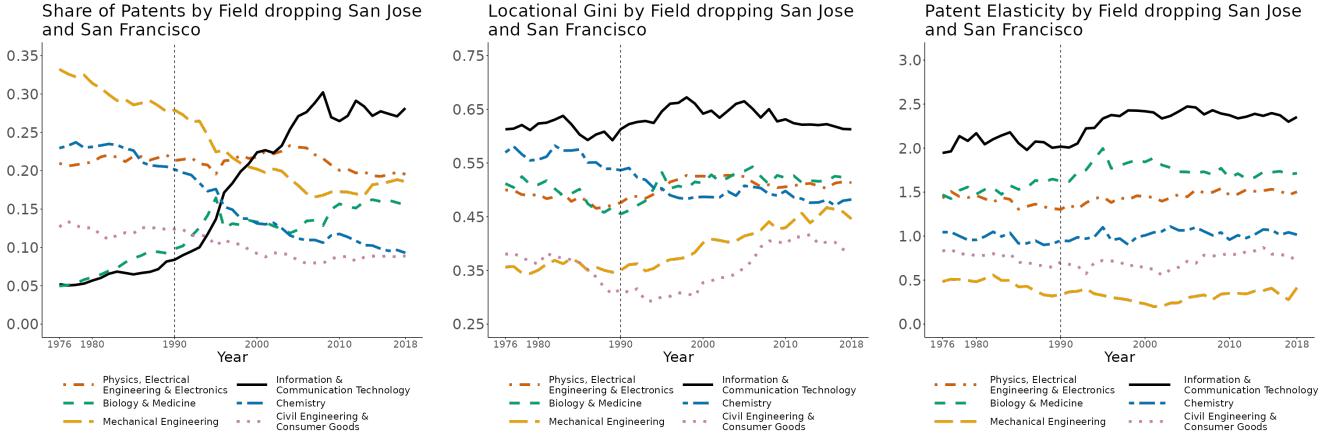


Figure 20: Trends in the share of annual patents (left panel), locational Gini coefficient of patents per capita (middle panel), and patent elasticity (right panel) by technology field dropping San Jose and San Francisco.

Figure 21 presents trends from a formal field-level decomposition of the locational Gini coefficient of patents per capita, excluding Silicon Valley. The left panel shows that without Silicon Valley, the post-1990 rise in the locational Gini is attenuated by 44%, and 67% of the remaining increase (compared to 52% in the full sample) driven by compositional changes across fields. The middle panel confirms that most of the cross-field composition is still explained by the rising annual share of patents in ICT. The right panel shows that the within-field contribution from ICT becomes smaller over time, turning negative after 2008, reflecting the decline in ICT spatial concentration once Silicon Valley is excluded. Since trends in patent elasticity by field remain largely unchanged after excluding Silicon Valley, the corresponding elasticity-based field decomposition is also unaffected.<sup>34</sup>

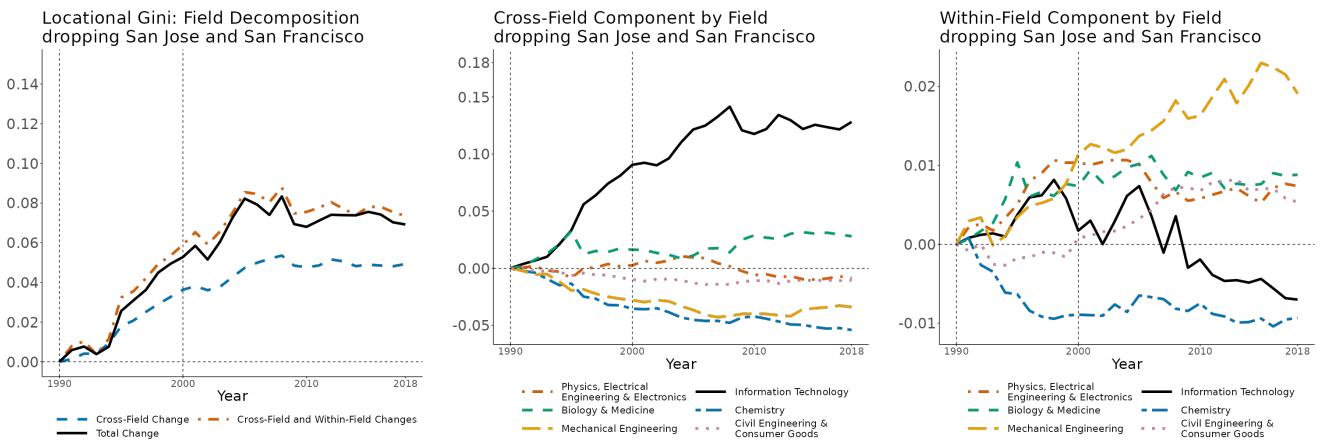


Figure 21: Trends in the decomposition of the locational Gini coefficient of patents per capita into cross-field and within-field components (left panel) and a further decomposition of the cross-field and within-field components into each individual field (middle and right panels respectively) after dropping San Jose and San Francisco.

<sup>34</sup>I omit the elasticity decomposition here, but results are available upon request.

### B.2.6 Trends and Decomposition by Subfield

To assess whether specific subfields drove the overall contribution of ICT to the rising spatial concentration of innovation in high-skill CZs, Figure 22 presents trends from the formal decomposition of the aggregate patent elasticity by subfield. The left panel shows that two-thirds of the post-1990 increase in aggregate patent elasticity is explained by compositional changes across subfields – compared to half in the field-level decomposition. The middle and right panels show the individual contributions of the top six subfields (ranked by their contribution to the post-1990 rise in patent elasticity) to the cross-subfield and within-subfield components, respectively. Most of the cross-subfield component is driven by the rising share of patents in *Digital Communication* and *Computer Technology*, which represent the core subfields in ICT. On the other hand, the top 5 subfields contribute similarly to the within-subfield component.

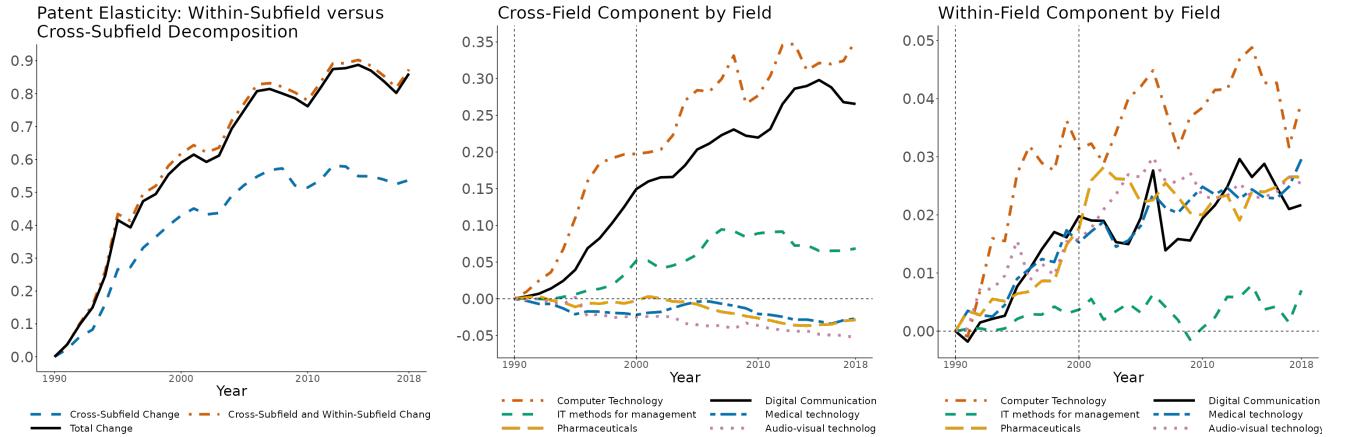


Figure 22: Trends in the subfield decomposition of the post-1990 increase in the elasticity of CZ patents per capita with respect to the 1990 college ratio. The left panel presents trends in the decomposition into cross-subfield and within-subfield components, while the middle and right panels display trends in the cross-subfield and within-subfield components respectively for the top six subfields (ranked by their contribution to the post-1990 rise in patent elasticity).

As discussed in Appendix B.2.3, a more complete picture of the spatial concentration of innovation requires examining the locational Gini coefficient of patents per capita. Figure 23 presents trends from the formal decomposition of the post-1990 increase in the locational Gini by subfield. The left panel shows that one-third of the rise in the aggregate locational Gini is driven by the colocation of subfields (i.e., the residual term). Given that the colocation component in the analogous field-level decomposition (left panel of Figure 18) is negligible, we can infer that this colocation occurred primarily between subfields within the same field. As in Figure 22, the middle and right panels of Figure 23 show the individual contributions of the six selected subfields to the cross-subfield and within-subfield components, respectively. Once again, most of the cross-subfield component is driven by the rising share of patents in *Digital Communication* and *Computer Technology*. Notably, the within-subfield contributions for these two areas begin to decline after 2005, suggesting a geographic shift in patenting activity in these two subfields from other centers of concentration toward high-skill CZs.

To assess both the level differences in spatial concentration and how trends evolved after 1990 relative to the

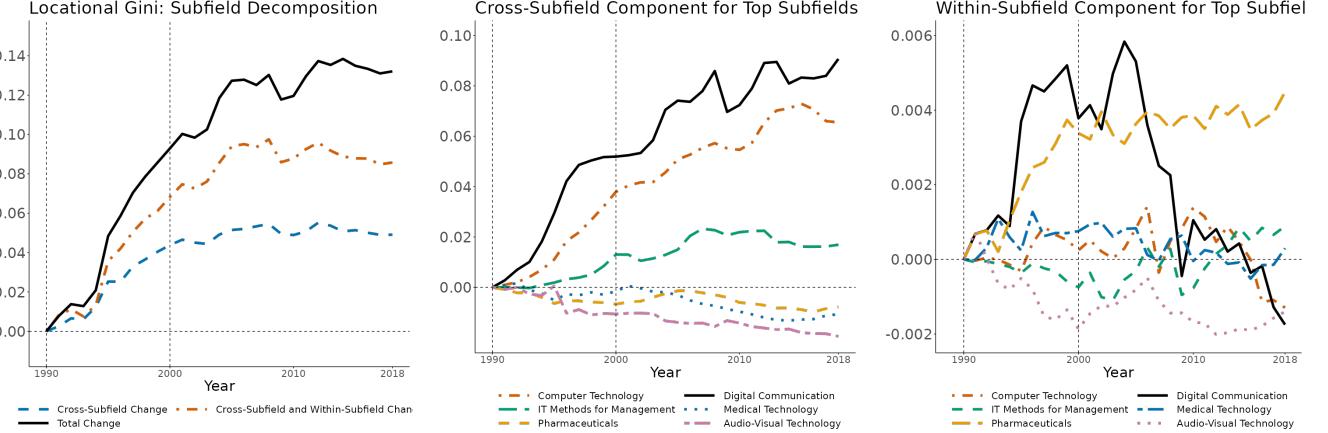


Figure 23: Trends in the subfield decomposition of the post-1990 increase in the locational Gini coefficient of patents per capita. The left panel presents trends in the decomposition of the aggregate locational Gini into cross-subfield and within-subfield components, while the middle and right panels display trends in the cross-subfield and within-subfield components respectively for the six subfields selected above.

preceding period, Figure 24 presents trends from 1976 to 2018 in the share of annual patents, the locational Gini coefficient, and patent elasticity for the six selected subfields. The left panel confirms that the share of patents in *Digital Communication* and *Computer Technology* increased dramatically after 1990. The middle panel shows that patents across all six subfields became more concentrated in high-skill CZs after 1990. Consistent with Figure 23, the right panel reveals a decline in the locational Gini for several of these subfields, suggesting a geographic shift in patenting activity away from other previously concentrated areas toward high-skill CZs.

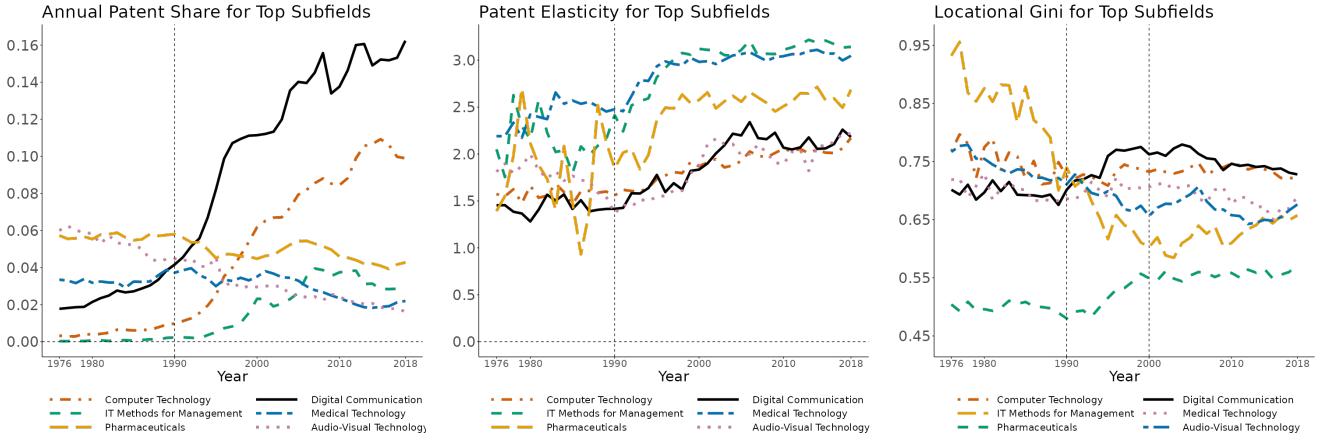


Figure 24: Trends for the top subfields by contribution to the post-1990 increase in overall patent elasticity: annual patent share (left panel), locational Gini coefficient of patents per capita (middle panel), and patent elasticity (right panel).

Table 5 lists the top 10 subfields ranked by their contribution to the within-subfield component. For comparison, it also includes subfields that appear in the top 10 by contribution to the cross-subfield component. While *Digital Communication* and *Computer Technology* exert a disproportionate influence on the cross-

Subfield	Field	Ranking	Elasticity Change Within Subfield	Ranking	Elasticity Change Across Subfield	Gini Change Within Subfield	Gini Change Across Subfield
Computer technology	2	1	0.0394	1	<b>0.3519</b>	-0.0017	0.0907
Audio-visual technology	1	2	0.0368	21	-0.0089	0.0034	-0.0014
Electrical machinery, apparatus, energy	1	3	0.0320	20	-0.0081	0.0026	-0.0033
Measurement	1	4	0.0296	33	-0.0270	0.0045	-0.0078
Optics	1	5	0.0265	34	-0.0287	0.0003	-0.0107
Organic fine chemistry	4	6	0.0253	35	-0.0532	-0.0014	-0.0194
Digital communication	2	7	0.0217	2	<b>0.2655</b>	-0.0013	0.0655
Medical technology	3	8	0.0195	4	0.0413	0.0080	0.0142
Telecommunications	2	9	0.0167	9	0.0087	0.0000	0.0032
Transport	5	10	0.0151	10	0.0029	0.0057	0.0022
Control	1	12	0.0096	7	0.0180	0.0012	0.0066
IT methods for management	2	15	0.0070	3	0.0684	0.0009	0.0169
Biotechnology	3	17	0.0052	8	0.0134	0.0004	0.0035
Pharmaceuticals	3	20	0.0038	5	0.0320	-0.0004	0.0086
Semiconductors	1	34	-0.0067	6	0.0235	0.0021	0.0069

Table 5: List of top 10 subfields by their contribution to the post-1990 increase in the elasticity within subfields, appended with the other subfields in the top 10 by contribution to the cross-subfield component.

subfield component (highlighted in bold), the within-subfield contributions are more evenly distributed across a broader set of subfields.

### B.3 Firm Decomposition of the Rising Concentration of Innovation in High-Skill CZs

Here I present additional results on the firm-level decompositions (**Facts 4 and 5** in the main text).

#### B.3.1 Firm Decomposition of the Post-1990 Compositional Shift in Innovation toward ICT

I begin by examining whether the post-1990 rise in the aggregate ICT patent share,  $s_t$ , reflects increases in ICT intensity within firms or instead the growing weight of firms that predominantly produce ICT patents. To appropriately account for firm entry and exit, I implement a modified Foster-Haltiwanger-Krizan (FHK) decomposition (Foster et al., 2001)<sup>35</sup>:

$$s_{t^*} - s_{1990} = \sum_{t=1991}^{t=t^*} \Delta s_t = \sum_{t=1991}^{t=t^*} \left[ \underbrace{\sum_{f \in C} \bar{w}_{f,t} \Delta s_{f,t}}_{\text{within-firm}} + \underbrace{\sum_{f \in C} (\bar{s}_{f,t} - \bar{S}_t) \Delta w_{f,t}}_{\text{between-firm}} + \underbrace{\sum_{f \in E} w_{f,t} (s_{f,t} - \bar{S}_t)}_{\text{entry}} \right. \\ \left. + \underbrace{\sum_{f \in X} w_{f,t-1} (s_{f,t-1} - \bar{s}_t)}_{\text{exit}} + \underbrace{\Delta (s_t - \bar{S}_t)}_{\text{residual}} \right], \quad (47)$$

where  $w_{f,t}$  is firm  $f$ 's patent weight (its share of total patents in year  $t$ ),  $s_{f,t}$  is the share of firm  $f$ 's patents in ICT, and  $S_t = \sum_f w_{f,t} s_{f,t}$  is the weighted average of firm-level ICT shares in year  $t$ . For any variable  $x$ ,  $\bar{x}_t \equiv \frac{x_t + x_{t+1}}{2}$  and  $\Delta x_t \equiv x_t - x_{t-1}$ . The sets  $C$ ,  $E$ , and  $X$  denote continuing, entering, and exiting firms, respectively. The five terms capture: (i) changes in ICT shares within firms, (ii) compositional changes among incumbents, (iii) entry, (iv) exit, and (v) a residual reflecting patents without assignees. The left panel of Figure 4 in Section 2.3.2 of the main text plots this decomposition and shows that the post-1990 increase in the aggregate ICT patent share is explained almost entirely by the between-firm component, i.e. the rising patent shares of ICT-oriented firms.

To validate the firm-level decomposition, I conduct two complementary group-level Shapley decompositions. First, I classify firms into two groups – ICT-oriented and non-ICT-oriented – based on their long-run ICT intensity. For each firm, I compute the weighted average of its annual ICT patent shares, using as weights the firm's relative national patent share across years. Firms with an average ICT share above the median are designated as ICT-oriented. I then implement the following group-level Shapley decomposition as follows:

$$s_{t^*} - s_{1990} = \sum_{t=1991}^{t=t^*} \Delta s_t = \sum_{t=1991}^{t=t^*} \left[ \underbrace{\sum_f \bar{w}_{g,t} \Delta s_{g,t}}_{\text{within-group}} + \underbrace{\sum_g \bar{s}_{g,t} \Delta w_{g,t}}_{\text{between-group}} \right] + \underbrace{\Delta \left( s_t - \sum_g s_{g,t} w_{g,t} \right)}_{\text{residual}}, \quad (48)$$

where  $w_{g,t}$  is group  $g$ 's patent weight (its share of total patents in year  $t$ ) and  $s_{g,t}$  is the share of group  $g$ 's

---

<sup>35</sup>The Shapley decomposition does not allow for an apples-to-apples comparison of the entry and exit terms, which comprise the elasticity *level*, with the between and within terms, which are in terms of changes.

patents in ICT. The left panel of Figure 25 shows that, even with this coarse two-group classification, the rising patent share of ICT-oriented firms (dot-dashed blue line) accounts for more than half of the post-1990 increase in the aggregate ICT patent share (solid black line).<sup>36</sup>

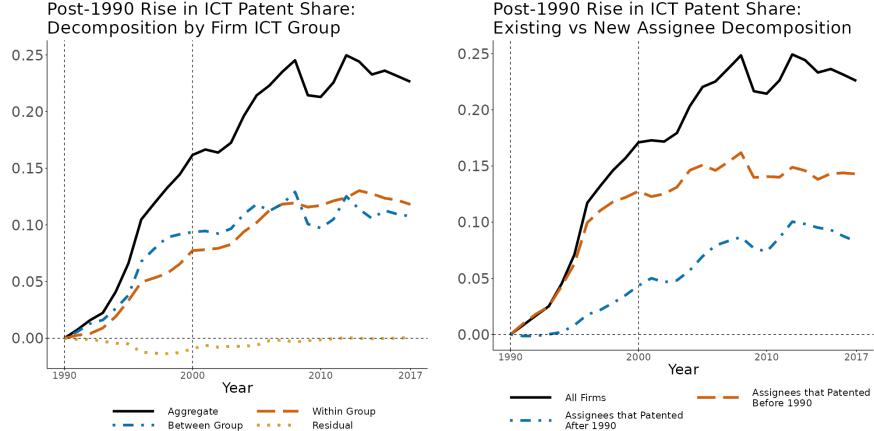


Figure 25: Decomposition of the post-1990 rise in the aggregate ICT patent share by firm long-run ICT intensity group (left) and firm entry cohort group (right).

Second, I classify firms by entry cohort – those that began patenting before 1990 versus after – and apply the same Shapley framework. I decompose the post-1990 increase in the aggregate ICT share into (i) growth in ICT shares among pre-1990 firms; (ii) growth in ICT shares among post-1990 entrants; and (iii) compositional changes due to the rising patent weight of post-1990 entrants. I attribute (ii) and (iii) jointly to post-1990 entrants. The right panel of Figure 25 shows that about three-quarters of the post-1990 increase in the aggregate ICT patent share is driven by firms that began patenting before 1990.

Taken together, both group-level decompositions confirm the firm-level FHK decomposition results: the post-1990 rise in the aggregate ICT patent share was mostly driven by the growing weight of existing ICT-oriented firms, rather than by within-firm changes or new entrants.

### B.3.2 Firm Decomposition of the Differences in Spatial Concentration between ICT and Non-ICT Innovation for each year

I next examine whether the greater concentration of ICT innovation in high-skill CZs relative to non-ICT, is driven by within-firms or between-firm differences. I restrict the sample to assignee-years where both ICT and non-ICT patent elasticities are well defined (i.e., the assignee has ICT and non-ICT patents respectively in at least 2 CZs in that year), covering 52% of all patents, and use the firm classification by long-run ICT intensity introduced above. The middle panel of Figure 4 in Section 2.3.2 displays the weighted average of assignee-level ICT and non-ICT patent elasticity for ICT-oriented versus non-ICT-oriented firms.<sup>37</sup> Notably, non-ICT patents produced by ICT-oriented firms are more concentrated in high-skill CZs than ICT patents

<sup>36</sup>I also consider an alternative firm classification by ICT intensity. Using annual ICT and non-ICT patent weights for each firm, I define three groups: (i) firms consistently ICT-oriented across all years, (ii) firms consistently non-ICT-oriented, and (iii) mixed firms. Decomposition results are similar under this classification and available upon request.

<sup>37</sup>Trends on the median and raw average are similar and available upon request.

produced by non-ICT-oriented firms. To formalize this comparison, I apply a Shapley decomposition of the aggregate ICT–non-ICT elasticity gap by year:

$$\alpha_{\text{ICT},t} - \alpha_{\text{nonICT},t} = \underbrace{\sum_f \bar{w}_{f,t} (\alpha_{f,\text{ICT},t} - \alpha_{f,\text{nonICT},t})}_{\text{within-firm}} + \underbrace{\sum_f \bar{\alpha}_{f,t} (w_{f,\text{ICT},t} - w_{f,\text{nonICT},t})}_{\text{between-firm}} + \underbrace{\varepsilon}_{\text{residual}} \quad (49)$$

where  $w_{f,k,t}$  denotes firm  $f$ 's patent weight in sector  $k = \{\text{ICT, non-ICT}\}$  in year  $t$  and  $\alpha_{f,k,t}$  represents the corresponding elasticity. The first term captures within-firm differences, the second term captures between-firm differences, and the residual reflects patents from firms with undefined elasticities. The right panel of Figure 4 shows that the aggregate ICT–non-ICT elasticity gap is driven more by the between-firm component, reflecting systematic cross-firm differences in patent geography rather than within-firm differences between ICT and non-ICT patents.

Several caveats apply. First, some of the firm-level elasticities are estimated with large standard errors, and outliers may influence the decomposition results. Second, almost half of all patents fall into assignee-years with undefined elasticities and are absorbed into the residual. To address these concerns, I estimate annual group-level elasticities for ICT-oriented versus non-ICT-oriented firms. The left panel of Figure 26 shows that these group-level elasticities closely mirror the assignee-weighted averages from Figure 4. I then apply a group-level Shapley decomposition of the aggregate ICT–non-ICT elasticity gap. The middle panel of Figure 26 confirms that between-group differences dominate within-group differences, while the residual is far smaller than in the firm-level decomposition. Together, these results confirm that cross-firm differences in patent geography, rather than within-firm differences, account for the bulk of the ICT–non-ICT elasticity gap.

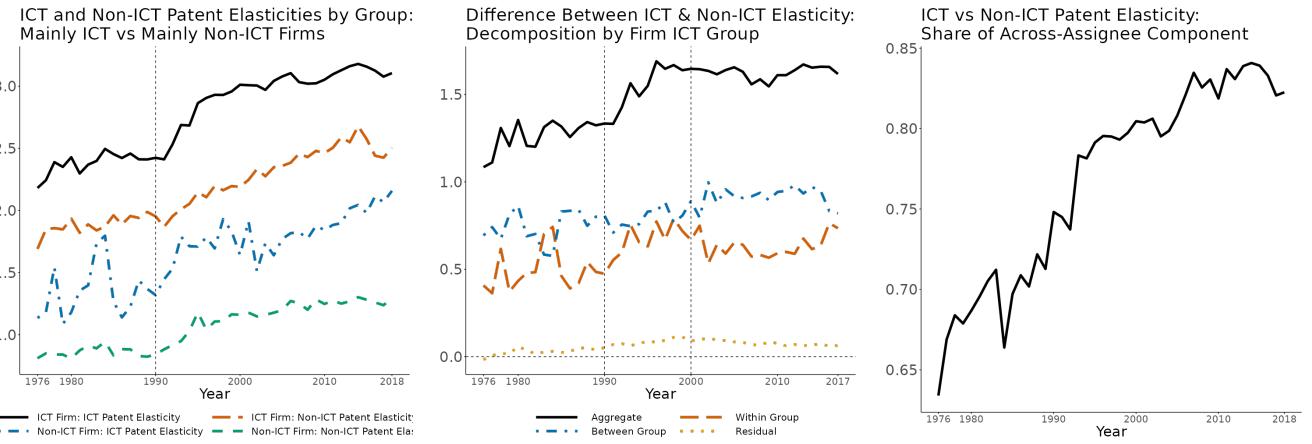


Figure 26: Group-level decompositions of the aggregate differences in the ICT and non-ICT patent elasticity for each year. The left panel plots trends in the group-level ICT and non-ICT patent elasticities for ICT-oriented versus non-ICT-oriented firms. The middle panel plots trends in the group decomposition by firm ICT intensity. The right panel plots the trends in the group decomposition by firm entry cohort.

A remaining question is whether the cross-firm differences in patent geography are driven primarily by pre-1990 firms or by post-1990 entrants. To distinguish between these forces, I use the firm cohort classification

introduced above and apply a similar group-level Shapley decomposition. For each year from 1990, I decompose the aggregate ICT–non-ICT elasticity gap into three components: (i) differences among pre-1990 firms; (ii) differences among post-1990 entrants; and (iii) differences between the two cohorts. I attribute (ii) and (iii) jointly to post-1990 entrants. The right panel of Figure 26 shows that the bulk of the aggregate ICT–non-ICT elasticity gap in most years after 1990 is explained by pre-1990 firms, with post-1990 entrants playing only a limited role.

### B.3.3 Firm Decomposition of the Post-1990 Rise in the Spatial Concentration of ICT and Non-ICT Innovation

Finally, I examine whether the post-1990 rise in aggregate ICT and non-ICT patent elasticities reflects changes within firms or reallocation across firms. For each case, I restrict the sample to firm-years with defined firm-level elasticities (account for 48% of ICT patents and 52% of non-ICT patents) and implement a modified FHK decomposition, analogous to that used for the post-1990 increase in the aggregate ICT patent share:

$$\alpha_{k,t^*} - \alpha_{k,1990} = \sum_{t=1991}^{t=t^*} \Delta \alpha_{k,t} = \sum_{t=1991}^{t=t^*} \left[ \underbrace{\sum_{f \in C} \bar{w}_{f,k,t} \Delta \alpha_{f,k,t}}_{\text{within-firm}} + \underbrace{\sum_{f \in C} (\bar{\alpha}_{f,k,t} - \bar{A}_{k,t}) \Delta w_{f,k,t}}_{\text{between-firm}} + \underbrace{\sum_{f \in E} w_{f,k,t} (\alpha_{f,k,t} - \bar{A}_{k,t})}_{\text{entry}} \right. \\ \left. + \underbrace{\sum_{f \in X} w_{f,k,t-1} (\alpha_{f,k,t-1} - \bar{A}_{k,t})}_{\text{exit}} + \underbrace{\Delta(\alpha_{k,t} - A_{k,t})}_{\text{residual}} \right], \quad (50)$$

where  $\alpha_{k,t}$  is the aggregate patent elasticity,  $\alpha_{f,k,t}$  is firm  $f$ 's patent elasticity,  $A_{k,t} = \sum_f w_{f,k,t} \alpha_{f,k,t}$  is the weighted average of firm-level elasticities in sector  $k \in \{\text{ICT, non-ICT}\}$  and year  $t$ . Other notation follows equation 47. Figure 5 in Section 2.3.2 plots the results: the left panel shows that between-firm reallocation explains more two-thirds of the post-1990 rise in the aggregate ICT patent elasticity, while the right panel shows that it accounts for nearly all of the corresponding rise in the non-ICT elasticity. Together, these results underscore the dominant role of firm reallocation.

Nonetheless, the same caveats noted for the firm-level decomposition of the ICT–non-ICT elasticity gap apply here. To address these concerns, I conduct two complementary *group*-level Shapley decompositions. First, I decompose by firm elasticity quartiles. For each firm, I compute the weighted average of its annual ICT patent elasticities, using as weights of each year is the firm's share of the national ICT patents in each year relative to the sum of these shares across all years. I then assign firms to four quartiles based on these average elasticities, pool patents within each quartile-year, and implement the following group-level Shapley

decomposition:

$$\alpha_{k,t^*} - \alpha_{k,1990} = \sum_{t=1991}^{t=t^*} \Delta \alpha_t = \sum_{t=1991}^{t=t^*} \left[ \underbrace{\sum_g \bar{w}_{g,k,t} \Delta \alpha_{g,k,t}}_{\text{within-group}} + \underbrace{\sum_g \bar{\alpha}_{g,k,t} \Delta w_{k,t}}_{\text{between-group}} + \underbrace{\Delta \left( \alpha_{k,t} - \sum_g w_{g,k,t} \alpha_{g,k,t} \right)}_{\text{residual}} \right] \quad (51)$$

where  $g$  indexes the four quartiles of firms and other notation follows equation 50.<sup>38</sup>

Second, I decompose by firm entry cohort. Using the cohort classification introduced above, I apply a similar group-level Shapley framework to the post-1990 increase in the aggregate ICT and non-ICT elasticities. Specifically, I separate the rise into three components: (i) changes among pre-1990 firms, (ii) changes among post-1990 entrants, and (iii) compositional shifts reflecting the growing patent weight of post-1990 entrants. I attribute (ii) and (iii) jointly to post-1990 entrants.

Figure 27 reports results from the two group-level decompositions of the post-1990 increase in the aggregate ICT elasticity. The left panel shows that most of the rise reflects the growing share of patents from higher-elasticity firms, rather than increases in elasticity within groups.<sup>39</sup> The right panel shows that firms patenting in ICT before 1990 account for about two-thirds of the rise in ICT elasticity during the 1990s and about half thereafter. Thus, the growth of incumbent ICT-patenting firms, rather than entry of new firms, drove the post-1990 increase in aggregate ICT patent elasticity.

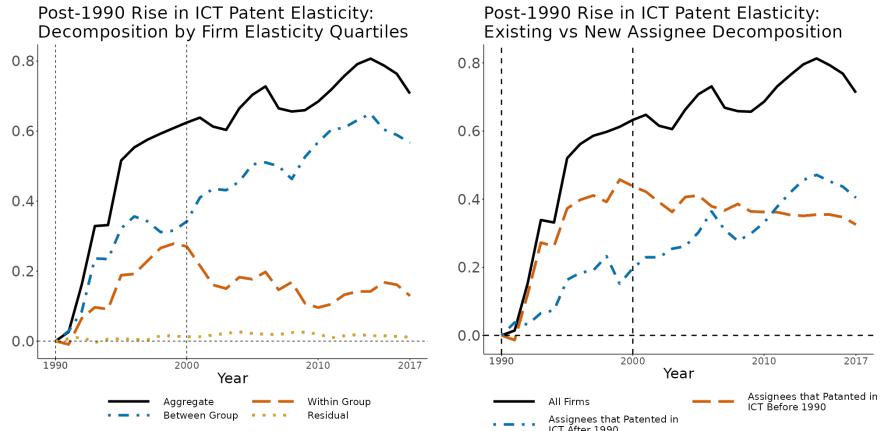


Figure 27: Group-level decomposition of the post-1990 rise in the aggregate ICT patent elasticity by firm ICT elasticity quartiles (left) and firm entry cohort (right).

Figure 28 shows the analogous group-level decompositions of the post-1990 increase in aggregate non-ICT elasticity. The left panel indicates that most of the rise reflects compositional shifts toward firms in higher-elasticity groups, while the right panel shows that incumbent firms with non-ICT patents prior to 1990

<sup>38</sup>In principle, there is an additional group of firms with undefined ICT patent elasticities because they patent in only one CZ in each year they patent. This group accounts for a small share of patents, and dropping it does not change the decomposition results. Decomposition results excluding this group are available on request.

<sup>39</sup>The within-group component accounts for just under half of the aggregate increase before 2000, in contrast to the much smaller within-firm component in the FHK decomposition. This reflects reallocation of patents across firms within each group.

account for the bulk of the increase.

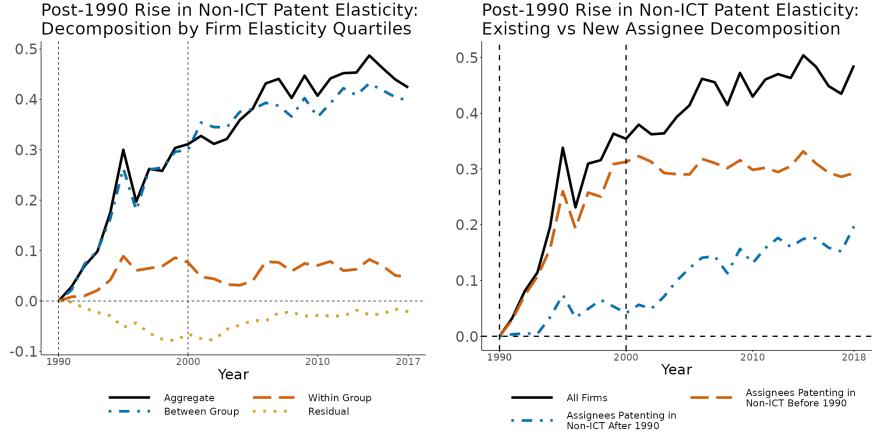


Figure 28: Group-level decomposition of the post-1990 rise in the aggregate non-ICT patent elasticity by firm non-ICT elasticity quartiles (left) and firm entry cohort (right).

These group-level decomposition results confirm that the post-1990 increase in both ICT and non-ICT elasticities is driven mainly by the rising patent shares of firms already concentrated in high-skill CZs. A key remaining question is whether this cross-firm reallocation reflects primarily the growing importance of ICT-oriented firms, or instead reallocation toward higher-elasticity firms within both ICT-oriented and non-ICT-oriented groups. To distinguish between both explanations, I first decompose the post-1990 rise in the aggregate ICT/non-ICT patent elasticity into the contributions of ICT-oriented versus non-ICT-oriented firms. In Figure 29, we see that most of the post-1990 rise in the non-ICT patent elasticity is driven by non-ICT-oriented firms (left panel), while that for the ICT patent elasticity is driven by ICT-oriented firms (middle panel). The right panel reports results from an FHK decomposition of the post-1990 increase in the non-ICT patent elasticity, using a sample restricted of non-ICT-oriented firms. These results show reallocation toward higher-elasticity firms within both the ICT- and non-ICT-oriented groups.

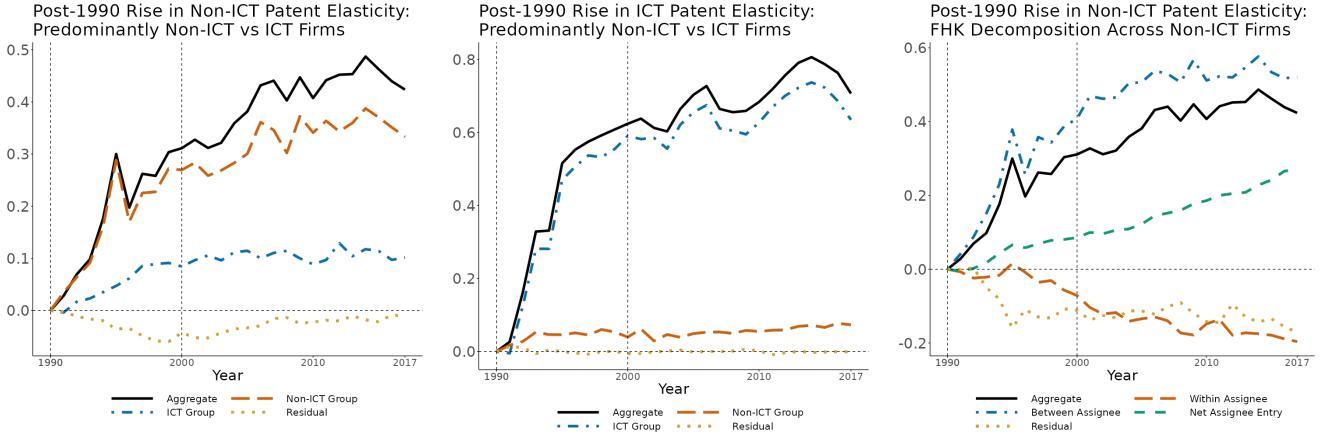


Figure 29: The left and middle panels present decompositions of the post-1990 rise in non-ICT and ICT patent elasticity respectively into contributions from the ICT- and non-ICT oriented firm groups. The right panel presents an FHK decomposition of the post-1990 rise in non-ICT elasticity using a sample restricted to non-ICT-oriented firms

## B.4 Details of the NSFNET

The National Science Foundation Network (NSFNET), established by the National Science Foundation and developed between 1986 and 1995, was created to facilitate collaboration among researchers at universities and military bases across the United States. It eventually became a crucial bridge between ARPANET – the first public packet-switched computer network operated by the Defense Advanced Research Projects Agency from 1969 to 1989 – and the commercial networks that began providing internet access to the public, particularly during the internet boom starting in 1995.

The history of the NSFNET is defined by five key events:

- Initial establishment: In 1985, the NSF funded the creation of five supercomputing centers. In 1986, it established a long-haul backbone network with a data speed of 56 Kbps, connecting these new centers to the existing supercomputing facility at the National Center for Atmospheric Research.
- First round of upgrading: On June 15, 1987, the NSF issued a solicitation to upgrade and expand the backbone network, addressing the overwhelming demand that had saturated the existing infrastructure. On November 24, 1987, a contract was awarded to a team comprising IBM, MCI, Merit, and the University of Michigan. By July 1988, the upgraded T1 backbone was completed, increasing the number of backbone sites from 6 to 13 and raising network speeds to 1.5 Mbps. The upgraded backbone also enabled connections to regional and campus networks. Each partner played a specific role: Merit developed user support and information services, IBM provided hardware, software, and network management tools, and MCI supplied transmission circuits with reduced tariffs.
- Second round of upgrading: In 1991, the NSF completed the upgrade from the T1 network to a T3 network, increasing the broadband speed to 45 Mbps and adding three new backbone sites.
- **Privatization:** Discussion about privatization began quietly in 1989 and became public by 1990. **In March 1991, the Acceptable Use Policy – which previously required the network to be used solely for research and education – was revised to allow private users.** In 1992, the NSF announced plans to decommission the NSFNET by 1995. By May 1993, it issued a solicitation to encourage more companies to contribute to the development of the Internet's privatized structure<sup>40</sup>
- Decommission: 1995

The most significant event in the history of the NSFNET for U.S. patenting – predominantly conducted by private firms – was the modification of the Acceptable Use Policy in March 1991. This change granted firms access to the NSFNET, significantly reducing their communication costs. Table 6 lists the NSFNET nodes, including the year each was established and the commuting (CZ) in which it is located. Note that although the NSFNET backbone consisted of just 11 nodes in 1991 and 15 in its full version in 1993, many of these

---

<sup>40</sup>Taken from Kesan, Jay P., and Rajiv C. Shah. 2001. "Fool Us Once, Shame on You—Fool Us Twice, Shame on Us: What We Can Learn from the Privatizations of the Internet Backbone Network and the Domain Name System." Washington University Law Quarterly 79: 89–220.

nodes connected to regional networks, thereby extending high-speed internet access to universities and firms across most US regions.

S/N	NSFNET Node Location	CZ	CZ Name	Year	Selected Universities
1	John von Neumann Center in Princeton, NJ	19600	Newark	1986	Princeton University
2	Cornell Theory Center in Ithaca, NY	18100	Elmira	1986	Cornell University
3	Pittsburgh Supercomputing Center in Pittsburgh, PA	16300	Pittsburgh	1986	Carnegie Mellon University, University of Pittsburgh
4	San Diego Supercomputer Center in San Diego, CA	38000	San Diego	1986	UC San Diego
5	National Center for Supercomputing Applications in Urbana, IL	23500	Decatur	1986	University of Illinois, Urbana-Champaign
6	National Center for Atmospheric Research in Boulder, CO	28900	Denver	1986	University of Colorado Boulder, Colorado School of Mines
7	Palo Alto, CA	37500	San Jose	1988	Stanford University
8	Houston, TX	32000	Houston	1988	Rice University
9	Ann Arbor, MI	11600	Detroit	1988	University of Michigan at Ann Arbor
10	College Park, Maryland	11304	Washington DC	1988	Georgetown University, University of Maryland
11	Salt Lake City, UT	36100	Salt Lake City	1988	
12	Seattle, WA	39400	Seattle	1988	
13	Lincoln, NE	28101	Lincoln	1988	
14	Cambridge, MA	20500	Boston	1991	Harvard University, MIT, Tufts University
15	Argonne National Laboratory in Lemont, IL	24300	Chicago	1991	University of Chicago, Northwestern University
16	Atlanta, Georgia	9100	Altanta	1991	Emory University

Table 6: List of NSFNET backbone nodes

## C Proofs and Extensions of Propositions and Lemmas

Lemmas 1 to 6 provide the equilibrium conditions that govern endogenous innovation and technology diffusion; Lemma 7 details the equilibrium conditions of dynamic worker mobility with frictions; Proposition 1 characterizes the *spatial direction of innovation* and; Propositions 2 to 4 *decomposes the welfare impacts of any shock to economic fundamentals into its transitory and long-run growth components*.

Specifically, in the proofs of Lemmas 1 and 2, I leverage techniques from Eaton and Kortum (2024) to provide a more general derivation of the multivariate productivity distribution, minimum cost distribution, and trade shares relative to Lind and Ramondo (2024). This generalization allows me to introduce imperfect competition, characterized by the price index (Lemma 3) and markup distribution (Lemma 4). Along with worker mobility (Lemma 7), I thus *introduce directed innovation in a dynamic spatial model*, characterized by the ratio of inventor wages across regions (Proposition 1) which incorporate the expected value of individual ideas (Lemma 5).

### C.1 Lemma 1 (Productivity Distribution)

*Proof.* I first derive the conditional distribution of applicabilities of an idea, then the goods productivity distribution across all ideas.

#### (i) Distribution of Applicabilities of an Idea

**Microfoundations** After an idea is discovered in region  $r$  at time  $t^*$ , independent stochastic processes of rate  $\frac{\tilde{\Omega}}{t-t^*} \cdot \Omega'_{ro,t^*}(t-t^*)dt$  govern its diffusion to each of the other regions  $o \neq r$ .  $\Omega'_{ro,t^*}(t-t^*)$  captures how the density or fertility of an idea in creating applications evolves with the age of the idea,  $t - t^*$ .

At each instance the idea arrives in region  $o$ , its applicability  $a$  is stochastic and drawn from a Pareto distribution:

$$H^A(a) = 1 - \left(\frac{a}{\underline{a}}\right)^{-\sigma}, \quad a \geq \underline{a}, \quad \sigma > \theta, \quad \underline{a} = \tilde{a} \cdot \Gamma\left(1 - \frac{\theta}{\sigma}\right)^{-\frac{1}{\theta}}.$$

Thus, the probability that each application is greater than  $a$  is:

$$p_a = 1 - H^A(a) = \Gamma\left(1 - \frac{\theta}{\sigma}\right)^{-\frac{\sigma}{\theta}} a^{-\sigma}.$$

The number of applications of this idea in region  $o$  that is greater than  $a$  by time  $t$  follows a binomial distribution with expectation:

$$\mathbb{E} = \Omega_{ro,t^*}(t-t^*) \cdot \Gamma\left(1 - \frac{\theta}{\sigma}\right)^{-\frac{\sigma}{\theta}} \cdot a^{-\sigma} \cdot \tilde{\Omega} \tilde{a}^\sigma.$$

I assume that as the number of applications get arbitrarily large ( $\tilde{\Omega} \rightarrow \infty$ ) and the applicability of a typical idea becomes arbitrarily small ( $\tilde{a}^\sigma \rightarrow 0$ ), its product converges to 1 ( $\tilde{\Omega} \cdot \tilde{a}^\sigma \rightarrow 1$ ). In this limit, the number

of applications greater than  $a$  by time  $t$  converges to a Poisson distribution with the same expectation. Thus, the number of applications with applicability  $A > a$  that have arrived in region  $o$  by time  $t$  of an idea discovered in region  $r$  at time  $t^*$  is distributed Poisson with parameter:

$$\lambda^A(a) = M_{ro,tt^*} a^{-\sigma}, \quad M_{ro,tt^*} \equiv \Omega_{ro,t^*}(t - t^*) \cdot \Gamma\left(1 - \frac{\theta}{\sigma}\right)^{-\frac{\sigma}{\theta}}, \quad a > 0,$$

which corresponds to equation (4) in the main text. The expected number of applications of this idea with applicability above  $a$  to ever arrive in region  $o$  is  $\Gamma\left(1 - \frac{\theta}{\sigma}\right)^{-\frac{\sigma}{\theta}} a^{-\sigma}$ . The expected share of these applications that arrives by time  $t$  is  $\Omega_{ro,t^*}(t - t^*)$ .

**Distribution of Applicabilities** Using this Poisson distribution, the distribution of the best applicability of this idea in region  $o$  by time  $t$  is given by the probability that there is no application more efficient than  $a$ :

$$F_{ro,t^*t}^A(a) = \mathbb{P}[A_o \leq a] = e^{-M_{ro,tt^*} a^{-\sigma}} = \exp\left[-\Gamma\left(1 - \frac{\theta}{\sigma}\right)^{-\frac{\sigma}{\theta}} \Omega_{ro,t^*}(t - t^*) \cdot a^{-\sigma}\right], \quad (52)$$

which is a Frechet distribution. Across all regions, the joint distribution of the best application in time  $t$  of an idea discovered in region  $r$  at time  $t^*$  is:

$$\begin{aligned} F_{r,t^*t}^A(a_1, \dots, a_N) &= \mathbb{P}[A_1 \leq a_1, \dots, A_N \leq a_N] = e^{-\sum_o M_{ro,tt^*} a_o^{-\sigma}} \\ &= \exp\left[-\Gamma\left(1 - \frac{\theta}{\sigma}\right)^{-\frac{\sigma}{\theta}} \sum_o \Omega_{ro,t^*}(t - t^*) \cdot a_o^{-\sigma}\right], \end{aligned} \quad (53)$$

which is the product of independent Frechet distributions in each region  $o$ .

## (ii) Goods Productivity Distribution

**Microfoundations** In region  $r$  at time  $t^*$ , ideas are discovered from a stochastic process at rate  $\lambda_{r,t^*} dt$ . At discovery, each idea  $i$  has stochastic quality drawn from a Pareto distribution:

$$H^Q(q) = 1 - \left(\frac{q}{\underline{q}}\right)^{-\theta}, \quad q \geq \underline{q}, \quad \theta > 0,$$

and pertains to the production of a specific good drawn from the unit interval. I assume that as the number of ideas gets arbitrarily large ( $\int_0^{t^*} \lambda_{r,t} dt \rightarrow \infty$ ), such that their quality converges to 0 ( $\underline{q}^\theta \rightarrow 0$ ), their product  $\underline{q}^\theta \int_0^{t^*} \lambda_{r,t} dt$  remains constant. Following a similar argument for the arrival of idea applications above, the number of ideas with quality greater than  $q$  that have been discovered in region  $r$  by time  $t^*$  is thus distributed Poisson with parameter:

$$\lambda_{r,t}^Q(q) = q^{-\theta} \underline{q}^\theta \int_0^{t^*} \lambda_{r,t} dt,$$

which corresponds to equation (3) in the main text.

After an idea is discovered in region  $r$  at time  $t^*$ , its diffusion to each of the other regions  $o \neq r$  is described in part (i) above. The efficiency of an idea  $i$  to produce the corresponding good  $\nu$  in region  $o$  at time  $t$  is the product of its quality and best applicability in region  $o$  at time  $t$ :

$$z_i = q_i a_{i,o,t}^*, \quad (54)$$

where  $a_{i,o,t}^*$  is drawn from equation (52). Each region  $o$  produces each good  $\nu$  with the most efficient idea:

$$Z_{o,t} = \max_i \{q_i a_{i,o,t}^*\}. \quad (55)$$

**Marginal Productivity Distribution In Each Region** Using the definition of idea efficiency in equation 54, I derive the Poisson process of idea application arrivals in region  $o$  incorporating both idea quality on its discovery in region  $r$  and idea applicability on its arrival in region  $o$ . The expected number of ideas with efficiency  $Z > z$  in region  $o$  at time  $t$  is the total number of ideas produced across all regions with quality above  $q = \frac{z}{A_{o,t}}$ , where  $A_{o,t}$  is the best application of the idea in region  $o$  by time  $t$  and held fixed when computing  $q$ :

$$\begin{aligned} \lambda_{o,t}^Z(z) &= z^{-\theta} \sum_{r=1}^N \int_0^t \int_0^\infty A^\theta dF_{ro,tt^*}^A(a) \lambda_{r,t^*} dt^* \\ &= z^{-\theta} \sum_{r=1}^N \int_0^t \Gamma\left(1 - \frac{\theta}{\sigma}\right) \cdot \left[ \Gamma\left(1 - \frac{\theta}{\sigma}\right)^{-\frac{\sigma}{\theta}} \Omega_{ro,t^*}(t-t^*) \right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^* \\ &= z^{-\theta} \sum_{r=1}^N \int_0^t [\Omega_{ro,t^*}(t-t^*)]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^* \end{aligned} \quad (56)$$

where the second equality comes from the mean of the Frechet distribution  $F_{ro,tt^*}^A(a)$  in equation (52). Thus the total number of ideas with efficiency above  $z$  in region  $o$  at time  $t$  is distributed Poisson with parameter  $\lambda_{o,t}^Z(z)$ :

$$P[I = i] = e^{-\lambda_{o,t}^Z(z)} \frac{[\lambda_{o,t}^Z(z)]^i}{i!}.$$

The probability that the  $i$ 'th most efficient idea has efficiency below  $z$  is the probability that at most  $i-1$  ideas have efficiency greater than  $z$ :

$$F_{o,t}^{Z(i)}(z) = e^{-\lambda_{o,t}^Z(z)} \sum_{l=0}^{i-1} \frac{[\lambda_{o,t}^Z(z)]^l}{l!}.$$

The marginal productivity distribution in region  $o$  at time  $t$  is given by the distribution of the most efficient idea:

$$F_{o,t}^{Z(1)}(z) = \mathbb{P}[Z_{o,t}^{(1)} \leq z] = \exp[-\lambda_{o,t}^Z(z)] = \exp[-T_{o,t} z^{-\theta}]$$

where:

$$T_{o,t} = \sum_{r=1}^N \int_0^t [\Omega_{ro,t^*}(t-t^*)]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^*$$

which correspond to equations (8) and (9) in the main text.

**Joint Productivity Distribution** I now derive the Poisson process for idea application arrivals across all regions incorporating quality and applicability, and consequently the joint productivity distribution. The expected number of ideas with efficiency above  $z_o$  in at least one region at time  $t$  is equivalent to the expected number of ideas with quality  $q > \min_o \frac{z_o}{A_{o,t}}$ , where  $A_{o,t}$  is the best application of the idea in region  $o$  at time  $t$  and held fixed when computing  $q$ :

$$\begin{aligned} \lambda_t^Z(z_1, \dots, z_N) &= \sum_{r=1}^N \int_0^t \int_0^\infty \left[ \min_o \left\{ \frac{z_o}{A_{o,t}} \right\} \right]^{-\theta} dF_{r,tt^*}^A(a_1, \dots, a_N) \lambda_{r,t^*} dt^* \\ &= \sum_{r=1}^N \int_0^t \int_0^\infty \max_o \left\{ A_o^\theta z_o^{-\theta} \right\} dF_{r,tt^*}^A(a_1, \dots, a_N) \lambda_{r,t^*} dt^* \\ &= \sum_{r=1}^N \int_0^t \Gamma \left( 1 - \frac{\theta}{\sigma} \right) \cdot \left[ \Gamma \left( 1 - \frac{\theta}{\sigma} \right)^{-\frac{\sigma}{\theta}} \sum_{o=1}^N \Omega_{ro,t^*}(t-t^*) \cdot z_o^{-\sigma} \right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^* \\ &= \sum_{r=1}^N \int_0^t \left[ \sum_{o=1}^N \Omega_{ro,t^*}(t-t^*) \cdot z_o^{-\sigma} \right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^* \end{aligned} \tag{57}$$

where the second equality comes from solving the integral:

$$\int_0^\infty \max_o \left\{ A_o^\theta z_o^{-\theta} \right\} dF_{r,tt^*}^A(a_1, \dots, a_N) = \mathbb{E} \left[ \max_o \left\{ A_o^\theta Z_o^{-\theta} \right\} \right] \tag{58}$$

via three steps: (i) computing the distribution of the transformed variable  $B_o \equiv \frac{a_o}{z_o}$ :

$$\mathbb{P}(B_o \leq b) = \mathbb{P} \left( \frac{A_o}{z_o} \leq b \right) = \mathbb{P}(A_o \leq bz_o) = \exp \left[ -\Gamma \left( 1 - \frac{\theta}{\sigma} \right)^{-\frac{\sigma}{\theta}} \Omega_{ro,t^*}(t-t^*) \cdot b^{-\sigma} z_o^{-\sigma} \right],$$

(ii) computing the distribution of the maximum of the transformed variable across regions  $o$ ,  $\max_o B_o$ :

$$\mathbb{P} \left( \max_o B_o \leq b \right) = \exp \left[ -\Gamma \left( 1 - \frac{\theta}{\sigma} \right)^{-\frac{\sigma}{\theta}} \sum_o \Omega_{ro,t^*}(t-t^*) \cdot z_o^{-\sigma} \cdot b^{-\sigma} \right],$$

and (iii) computing the moment  $\mathbb{E} [\max_o B_o^\theta]$  using the known formula for Frechet random variables:

$$\mathbb{E} \left[ \max_o B_o^\theta \right] = \mathbb{E} \left[ \left( \max_o B_o \right)^\theta \right] = \Gamma \left( 1 - \frac{\theta}{\sigma} \right) \left[ \Gamma \left( 1 - \frac{\theta}{\sigma} \right)^{-\frac{\sigma}{\theta}} \sum_o \Omega_{ro,t^*}(t-t^*) \cdot z_o^{-\sigma} \right]^{\frac{\theta}{\sigma}}.$$

Using this Poisson process, the number of ideas with efficiency above  $z_o$  in at least one region by time  $t$  is distributed Poisson with parameter  $\lambda_t^Z(z_1, \dots, z_N)$ :

$$\mathbb{P}[I = i] = e^{-\lambda_t^Z(z_1, \dots, z_N)} \frac{[\lambda_t^Z(z_1, \dots, z_N)]^i}{i!}.$$

The probability that the  $i$ 'th most efficient idea has efficiency below  $z$  is the probability that at most  $i - 1$  ideas have efficiency greater than  $z$ :

$$F_t^{Z(i)}(z_1, \dots, z_N) = e^{-\lambda_t^Z(z_1, \dots, z_N)} \sum_{l=0}^{i-1} \frac{[\lambda_t^Z(z_1, \dots, z_N)]^l}{l!}.$$

The joint productivity distribution across all regions at time  $t$  is given by the distribution of the most efficient idea:

$$\mathbb{P}[Z_1 \leq z_1, \dots, Z_N \leq z_N] = F_t^{Z(1)}(z_1, \dots, z_N) = \exp \left[ - \sum_{r=1}^N \int_0^t \left[ \sum_{o=1}^N \Omega_{ro,t^*}(t - t^*) \cdot z_o^{-\sigma} \right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^* \right],$$

which corresponds to equation (7) in the main text.  $\square$

## C.2 Lemma 2 (Distribution of the Minimum Costs and Trade Shares)

*Proof.* I first derive the unconditional cost distribution, then the conditional cost distribution in each region, the joint cost distribution, and finally the equilibrium trade shares.

### (i) Distribution of the Minimum Cost in each Destination

Each destination sources goods from the cheapest location. Thus, what matters is the number of ideas whose goods can be delivered to region  $d$  below cost  $c$  by some region  $o$  by time  $t$ . The expected number of these ideas is the expected number of ideas with efficiency greater than  $z_o = \frac{w_{o,t}\tau_{od,t}}{c}$  in at least one region  $o$ :

$$\lambda_{d,t}^C(c) = \lambda_t^Z \left( \frac{w_{1,t}\tau_{1d,t}}{c}, \dots, \frac{w_{N,t}\tau_{Nd,t}}{c} \right) = c^\theta \sum_{r=1}^N \int_0^t \left[ \sum_{o=1}^N \Omega_{ro,t^*}(t - t^*) \cdot (w_{o,t}\tau_{od,t})^{-\sigma} \right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^*. \quad (59)$$

Thus, the number of ideas whose goods can be delivered to region  $d$  below cost  $c$  by time  $t$  – regardless of production and innovation location – is distributed Poisson with parameter  $\lambda_{d,t}^C(c)$ . Ranking these ideas by their costs in destination  $d$ ,  $C^{(1)} < C^{(2)} < C^{(3)} < \dots$ , the probability that the  $i$ 'th idea is above cost  $c$  is given by:

$$e^{-\lambda_{d,t}^C(c)} \sum_{l=0}^{i-1} \frac{[\lambda_{d,t}^C(c)]^l}{l!}$$

Thus the probability that the  $i$ 'th idea is below cost  $c$  is given by:

$$G_{d,t}^{(i)}(c) = 1 - e^{-\lambda_{d,t}^C(c)} \sum_{l=0}^{i-1} \frac{\left[\lambda_{d,t}^C(c)\right]^l}{l!} = 1 - \mathbb{P}\left[C_{1d,t}^{(i)} > c, \dots, C_{Nd,t}^{(i)} > c\right]. \quad (60)$$

In particular, the distribution of the lowest cost is given by:

$$G_{d,t}^{(1)}(c) = 1 - e^{-\lambda_{d,t}^C(c)} = 1 - \exp\left\{-\left[\sum_{r=1}^N \int_0^t \left[\sum_{o=1}^N \Omega_{ro,t}(t-t^*) (w_{o,t}\tau_{od,t})^{-\sigma}\right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^*\right] c^\theta\right\}.$$

## (ii) Equilibrium Trade Shares

Because different production locations may potentially produce goods from the same idea, the number of potential goods produced by each region  $o$  is not independent. This correlation across production locations implies that the straightforward way of computing trade shares does not hold,  $\pi_{od,t} \neq \frac{\lambda_{od,t}^C(c)}{\lambda_{d,t}^C(c)}$ . Instead, trade shares take on a nested form reflecting this correlation.

Since the number of ideas is independent across research locations, the share of goods purchased by destination  $d$  that were produced from ideas discovered in research location  $r$  is given by:

$$\phi_{rd,t} = \frac{\lambda_{rd,t}^C(c)}{\lambda_{d,t}^C(c)} = \frac{\int_0^t \left[\sum_{o=1}^N \Omega_{ro,t^*}(t-t^*) \cdot (w_{o,t}\tau_{od,t})^{-\sigma}\right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^*}{\sum_{r'=1}^N \int_0^t \left[\sum_{o=1}^N \Omega_{r'o,t^*}(t-t^*) \cdot (w_{o,t}\tau_{od,t})^{-\sigma}\right]^{\frac{\theta}{\sigma}} \lambda_{r',t^*} dt^*}.$$

For each idea discovered in region  $r$  at time  $t^*$  whose good is eventually sold to region  $d$  at time  $t$ , its quality  $q$  and cost  $c$  are fixed and the number of applications in each production location  $o$  is independent. Thus, the share of these ideas whose goods are produced in region  $o$  is given by the number of applications with applicability above  $a = \frac{w_{o,t}\tau_{od,t}}{q \cdot c}$  in region  $o$  divided by the total number of applications with applicability above  $a = \frac{w_{o',t}\tau_{o'd,t}}{q \cdot c}$  for all regions  $o'$ :

$$\varphi_{o|rd,t^*t} = \frac{\lambda_{o|rd,t}^A\left(\frac{w_{o,t}\tau_{od,t}}{q \cdot c}\right)}{\sum_{o'} \lambda_{o'|rd,t}^A\left(\frac{w_{o',t}\tau_{o'd,t}}{q \cdot c}\right)} = \frac{\Omega_{ro,t^*}(t-t^*) (w_{o,t}\tau_{od,t})^{-\sigma}}{\sum_{o'=1}^N \Omega_{r{o'},t^*}(t-t^*) (w_{o',t}\tau_{o'd,t})^{-\sigma}}.$$

where the function for the Poisson arrival of applications  $\lambda^A(a)$  above  $a$  comes from equation (4) in the main text. Since this share is independent of  $q$  and hence independent across ideas, the share of goods sold in destination  $d$  at time  $t$  whose ideas were discovered in region  $r$  and goods produced in region  $o$  (the trilateral

share) is given by:

$$\pi_{rod,t} = \int_0^t \frac{\left[ \sum_{o'=1}^N \Omega_{ro',t^*}(t-t^*) \cdot (w_{o',t}\tau_{o'd,t})^{-\sigma} \right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*}}{\sum_{r'=1}^N \int_0^t \left[ \sum_{o'=1}^N \Omega_{r'o',t^*}(t-t^*) \cdot (w_{o',t}\tau_{o'd,t})^{-\sigma} \right]^{\frac{\theta}{\sigma}} \lambda_{r',t^*} dt^*} \frac{\Omega_{ro,t^*}(t-t^*) (w_{o,t}\tau_{od,t})^{-\sigma}}{\sum_{o'=1}^N \Omega_{ro',t^*}(t-t^*) (w_{o',t}\tau_{o'd,t})^{-\sigma}} dt^*.$$

The trade shares are obtained by summing these trilateral shares across research locations:  $\pi_{od,t} = \sum_r \pi_{rod,t}$ . These expressions correspond to equations (12)-(15) in the main text.  $\square$

### C.3 Lemma 3 (Price Index)

*Proof.* Under Bertrand competition, consumers regard different varieties (i.e. different idea applications) of a good as perfect substitutes. Since all goods are essential for the production of the local final good, the price charged for each good is its second lowest cost, drawn from the following distribution:

$$G_{d,t}^{(2)}(c) = 1 - e^{-\lambda_{d,t}^C(c)} [1 + \lambda_{d,t}^C(c)] \quad (61)$$

where  $\lambda_{d,t}^C(c)$  comes from equation (59). Notice that this is a special case of equation (60) with  $i = 2$ . To simplify the subsequent computations, I define:

$$\tilde{\Phi}_{d,t} \equiv \sum_{r=1}^N \int_0^t \left[ \sum_{o=1}^N \Omega_{ro,t^*}(t-t^*) \cdot (w_{o,t}\tau_{od,t})^{-\sigma} \right]^{\frac{\theta}{\sigma}} \lambda_{r,t^*} dt^*, \quad (62)$$

so that  $\lambda_{d,t}^C(c) = \tilde{\Phi}_{d,t} c^\theta$  and  $G_{d,t}^{(2)}(c) = 1 - e^{-\tilde{\Phi}_{d,t} c^\theta} [1 + \tilde{\Phi}_{d,t} c^\theta]$ .

Aggregating prices across all varieties yields the price index in each region:

$$P_{d,t} = \exp \int_0^1 \ln p_{d,t}(\nu) d\nu = \exp \int_0^\infty \ln c \, dG_{d,t}^{(2)}(c) = \gamma \tilde{\Phi}_{d,t}^{-\frac{1}{\theta}}$$

where  $\gamma$  is the Euler-Mascheroni constant. This corresponds to equation (18) in the main text.  $\square$

### C.4 Lemma 4 (Markup Distribution)

*Proof.* Under Bertrand competition, the lowest cost producer of each good claims the entire market for that good, charging the highest markup that deters any competitor from entering. With the Cobb-Douglas aggregation across individual goods to produce the local sectoral final good (equation 10), the price of each good is given by the second lowest cost. Consequently, the markup of each good is the ratio of the lowest cost to the second lowest cost. I now apply the general derivation of the price gap distribution from Eaton and Kortum (2024) in my setting with idea applications, then use it to obtain the ratio of the lowest cost to the second lowest cost.

### (i) Distribution of Transformed Costs

To derive the price gap distribution, it is helpful to work with transformed costs  $U \equiv \tilde{\Phi}C^\theta$  where  $\tilde{\Phi}$  is given by equation (62). I first derive the distribution for  $U^{(i)}$ . Ranking the  $U^{(i)}$ 's like the  $C^{(i)}$ 's, the distribution of the lowest transformed cost is the probability that there is no transformed cost below  $u$ :

$$\mathbb{P}[U^{(1)} \leq u] = 1 - e^{-u},$$

which is the unit exponential distribution. Given a lower transformed cost  $U^{(i)} = u_i$ , the probability that the next lowest transformed cost  $U^{(i+1)}$  is less than  $u$  is simply the probability that there is no transformed cost between  $u_i$  and  $u$ :

$$\mathbb{P}[U^{(i+1)} \leq u | U^{(i)} = u_i] = 1 - e^{-(u-u_i)}. \quad (63)$$

Thus, the distribution of the gap conditional on the lower transformed cost is unit exponential. This implies that the set of ordered transformed costs can be drawn from independent unit exponential distributions  $V_i$  where:

$$U^{(1)} = V_1, \quad U^{(2)} = U^{(1)} + V_2, \quad U^{(3)} = U^{(2)} + V_3, \quad \dots \quad (64)$$

Since  $U^{(i)}$  is the sum of  $i$  independent unit exponential distributions, it follows an Erlang distribution with shape parameter  $i$  and rate parameter 1. Its probability density function is:

$$f^{(i)}(u) = \frac{u^{i-1}e^u}{(i-1)!} \quad (65)$$

and cumulative distribution function is:

$$F^{U^{(i)}}(u) = \mathbb{P}[U^{(i)} \leq u] = 1 - e^{-u} \sum_{l=0}^{i-1} \frac{u^l}{l!} = \int_0^u \frac{x^{i-1}e^x}{(i-1)!} dx. \quad (66)$$

Notice that the distribution of transformed costs can also be directly obtained from the cost distribution (equation 60). This derivation provides intuition on why the cost and transformed cost distributions are Erlang.

### (ii) Conditional Distribution of Lower Transformed Costs Given Higher Transformed Costs

Using the distribution of transformed costs  $U^{(i)}$ , I now derive the probability of the lower cost given a higher cost. The joint probability density function of  $U^{(i)}$  and  $U^{(i+1)}$  is given by:

$$\begin{aligned} f^{(i,i+1)}(u_i, u_{i+1}) &= \frac{d}{du_{i+1}} \left\{ \mathbb{P}[U^{(i+1)} \leq u_{i+1} | U^{(i)} = u_i] \cdot \mathbb{P}[U^{(i)} = u_i] \right\} \\ &= \frac{d}{du_{i+1}} \left\{ \left(1 - e^{-(u_{i+1}-u_i)}\right) \frac{u_i^{i-1}e^{u_i}}{(i-1)!} \right\} \\ &= \frac{u_i^{i-1}e^{u_{i+1}}}{(i-1)!} \end{aligned} \quad (67)$$

Thus the conditional distribution of the lower transformed costs (i.e. higher quality draw) is given by:

$$\mathbb{P} \left[ U^{(i)} \leq u | U^{(i+1)} = u_{i+1} \right] = \int_0^u \frac{f^{(i,i+1)}(x, u_{i+1})}{f^{(i+1)}(u_{i+1})} dx = \int_0^u \frac{x^{i-1} e^{u_{i+1}}}{(i-1)!} \frac{i!}{u_{i+1}^i e^{u_{i+1}}} dx = \left( \frac{u}{u_{i+1}} \right)^i \quad (68)$$

which is a Pareto distribution with shape parameter  $i$  and scale parameter  $u_{i+1}$ .

### (iii) Markup Distribution under Bertrand Competition

Using this result, the markup distribution conditional on the second lowest cost is given by:

$$\begin{aligned} G^{(2)/(1)}(m) &\equiv \mathbb{P} \left[ \frac{C^{(2)}}{C^{(1)}} \leq m | C^{(2)} = c_2 \right] = \mathbb{P} \left[ \frac{U^{(2)}}{U^{(1)}} \leq m^\theta | U^{(2)} = u_2 \right] \\ &= \mathbb{P} \left[ U^{(1)} \geq u_2 m^{-\theta} | U^{(2)} = u_2 \right] = 1 - \left( \frac{m^{-\theta} u_2}{u_2} \right) = 1 - m^{-\theta}. \end{aligned} \quad (69)$$

Remarkably, this distribution is independent of the second lowest cost  $c_2$ . Intuitively, this result arises because the conditional distribution of the higher quality draw is Pareto and hence scale invariant. That this, the ratio of the highest quality to the second highest quality idea is independent of the second highest quality idea.

Note that the markup distribution conditional on the lowest cost is different and given by:

$$\mathbb{P} \left[ \frac{C^{(2)}}{C^{(1)}} \leq m | C^{(1)} = c_1 \right] = \mathbb{P} \left[ U^{(1)} \leq u_1 m^\theta | U^{(1)} = u_1 \right] = 1 - \exp \left[ -\Phi c_1^\theta (m^{1/\theta} - 1) \right] \quad (70)$$

In contrast to equation (69), this distribution depends on the lowest cost. Intuitively, a higher value of the higher quality draw implies that the cutoff/second highest quality draw must have been higher.  $\square$

## C.5 Lemma 5 (Expected Value of an Idea)

*Proof.* Total profits earned at time  $s$  in region  $d$  is given by  $\Pi_{d,t}^k = \frac{X_{d,t}^k}{1+\theta}$ , as described in equation (20). The share of these profits from ideas discovered in region  $r$  at time  $t^*$  is  $\phi_{rd,t^*t}^k$ . Since the flow rate of ideas in region  $r$  at time  $t^*$  is  $\lambda_{r,t^*}^k$ , the expected flow of profits at time  $s$  in region  $d$  of an idea discovered in region  $r$  at time  $t^*$  is  $\frac{\phi_{rd,t^*t}^k}{\lambda_{r,t^*}^k} \frac{X_{d,t}^k}{1+\theta}$ . Accounting for changes in the purchasing power in region  $r$  over time and discounting future flows yields equation (21). Notice that  $\phi_{ld,t^*t}^k$  conditions on idea cohort and hence is a direct measure on whether the idea remains the lowest cost one in destination  $d$ .  $\square$

## C.6 Lemma 6 (Market Clearing)

*Proof.* The market clearing condition is slightly more complex relative to standard trade models due to the transfer of profits across regions. To account for these transfers, I distinguish across profits earned from innovation  $\bar{\Pi}$ , production  $\Pi^*$ , and sales  $\Pi$  in each region. Income in each region  $Y_{d,t}$  equals total expenditure

from the region, which comes from two sources: final spending by production workers and profits earned from innovation by firms. Thus the market clearing condition is given by:

$$\begin{aligned}
w_{o,t}^{k,G} L_{o,t}^{k,G} + \Pi_{o,t}^k * &= \sum_d \pi_{od,t}^k \ell^k \left[ \sum_k w_{d,t}^{k,G} L_{d,t}^{k,G} + \bar{\Pi}_{d,t}^k \right] \\
\implies \frac{1+\theta}{\theta} w_{o,t}^{k,G} L_{o,t}^{k,G} &= \sum_d \pi_{od,t}^k \ell^k \left[ \sum_k \left( w_{d,t}^{k,G} L_{d,t}^{k,G} + \sum_r \varphi_{dr,t}^k \Pi_{r,t}^k * \right) \right] \\
\implies \frac{1+\theta}{\theta} w_{o,t}^{k,G} L_{o,t}^{k,G} &= \sum_d \pi_{od,t}^k \ell^k \left[ \sum_k \left( w_{d,t}^{k,G} L_{d,t}^{k,G} + \sum_r \varphi_{dr,t}^k \frac{1}{\theta} w_{r,t}^{k,G} L_{r,t}^{k,G} \right) \right]
\end{aligned} \tag{71}$$

where  $\varphi_{dr,t}^k$  are the idea adoption shares and  $\pi_{od,t}^k$  are the trade shares. Notice that I eliminate the profit terms in the market clearing condition by exploiting the fact that profits earned from production is a constant multiple of income earned by production workers in the region-sector:

$$w_{o,t}^{k,G} L_{o,t}^{k,G} + \frac{1}{1+\theta} \sum_d \pi_{od,t}^k X_{d,t}^k = \sum_d \pi_{od,t}^k X_{d,t}^k.$$

This enhances the tractability and simplifies the potential quantification of the model.  $\square$

## C.7 Lemma 7 (Expected Worker Value and Worker Mobility Shares)

*Proof.* Note that the individual worker mobility problem, as defined by equations 25-26, represents a continuous time extension of Caliendo et al. (2019) (henceforth CDP) and incorporates switching between production and research alongside mobility across regions and sectors. In what follows, I demonstrate how the equilibrium mobility shares, as expressed in equation 28, can be derived under these extensions.

First, transitioning from discrete to continuous time with the possibility of worker movement complicates the distinction between current versus future payoffs. To overcome this challenge, I assume that there is a Poisson arrival process governing when *all* workers can move, with an arrival rate of 1, as described in Assumption 3 in the main text. This setup is a simplified version of Arcidiacono et al. (2016). With a Poisson arrival process of rate 1, the time until the next arrival at  $t'$  follows an exponential distribution with rate 1, where the probability density function is given by  $e^{-(t'-t)}$ . The present value at time  $t$  of the payoff at time  $t'$  is given by:  $V_{t'} e^{-\zeta(t'-t)}$ , where  $\zeta$  is the discount rate. Thus, the expected value at time  $t$  of the payoff at time  $t'$  is given by:

$$\text{Expected Payoff} = V_{t'} \cdot \int_0^\infty e^{-\zeta(t'-t)} e^{-(t'-t)} d(t' - t) = \frac{1}{1+\rho} V_{t'}.$$

This is why the individual worker mobility problem in equation 25 has the discount factor  $\frac{1}{1+\rho}$ . Note that my formulation is a strict generalization of CDP. Though my discount factor  $\frac{1}{1+\rho}$  is isomorphic to  $\beta$  in the migration problem in CDP and  $\mathbb{E}_t(t') = t + 1$ , the actual migration time  $t'$  is stochastic and depends on the exact realization of the Poisson process for move arrivals.

Second, I incorporate switching between production and research, where individual-specific idiosyncratic shocks for each market are drawn from a multivariate Gumbel or generalized extreme value distribution with symmetric correlation between production and research across all region-sectors. The functional form of this distribution is provided in equation 26, with the correlation function defined as:

$$\check{F}\left(\{\exp(-\epsilon_{o,t}^{s,n})\}_{o=1,\dots,N}^{s=\{\text{ICT,non-ICT}\},n=\{G,R\}}\right) = \sum_o \sum_s \left( \sum_n \exp(-\epsilon_{o,t}^{s,n})^{\frac{\Upsilon}{v}} \right)^v, \quad (72)$$

where  $v$  captures the correlation between production and research across region-sectors, and  $\Upsilon$  is the scale parameter that adjusts the sensitivity of the expected value in production or research to differences in the deterministic components of value in these activities. Note that this correlation function is homogeneous of degree  $\log \Upsilon$  (instead of 1), similar to Ben-Akiva and Francois (1983), but otherwise retains the same properties listed in McFadden (1978).

I now use this correlation function to derive the option value of moving and the resulting mobility shares. Since the option value of moving is defined as:

$$\Phi_{d,t}^{k,h} = \mathbb{E}_\epsilon \left[ \max_{o,s,n} \left\{ \frac{1}{1+\rho} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn} + \epsilon_{o,t}^{s,n} \right\} \right],$$

I first derive the distribution of the maximum of random variables whose joint distribution follows the GEV distribution in equation (26). Let  $U_{o,t}^{s,n} = \frac{1}{1+\rho} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn} + \epsilon_{o,t}^{s,n}$  and  $U = \max_{o,s,n} U_{o,t}^{s,n}$ . Extending Proposition 1 in Choi and Moon (1997), the cumulative distribution function is:

$$\begin{aligned} F_U(u) &\equiv \mathbb{P}(U \leq u) \\ &= \mathbb{P}(U_1^{\text{ICT},G} \leq u, \dots, U_N^{\text{non-ICT},R} \leq u) \\ &= \exp \left\{ -\check{F} \left( \exp \left[ \frac{1}{1+\rho} V_{1,t'}^{\text{ICT},G} - \kappa_{d1,t}^{k\text{ICT},hG} - u \right], \dots, \exp \left[ \frac{1}{1+\rho} V_{1,t'}^{\text{non-ICT},R} - \kappa_{dN,t}^{k\text{non-ICT},hR} - u \right] \right) \right\} \\ &= \exp \left\{ -\exp(-\Upsilon u) \check{F} \left( \exp \left[ \frac{1}{1+\rho} V_{1,t'}^{\text{ICT},G} - \kappa_{d1,t}^{k\text{ICT},hG} \right], \dots, \exp \left[ \frac{1}{1+\rho} V_{1,t'}^{\text{non-ICT},R} - \kappa_{dN,t}^{k\text{non-ICT},hR} \right] \right) \right\} \\ &= \exp \left\{ -\exp \left( -\Upsilon \left[ u - \frac{1}{\Upsilon} \log \left[ \check{F} \left( \exp \left[ \frac{1}{1+\rho} V_{1,t'}^{\text{ICT},G} - \kappa_{d1,t}^{k\text{ICT},hG} \right], \dots, \exp \left[ \frac{1}{1+\rho} V_{1,t'}^{\text{non-ICT},R} - \kappa_{dN,t}^{k\text{non-ICT},hR} \right] \right) \right] \right] \right) \right\} \end{aligned}$$

where the second last equality comes from the correlation function being homogeneous of degree  $\log \Upsilon$ . Thus, the distribution of the maximum is a Gumbel distribution with scale parameter  $\Upsilon$  and location parameter  $\frac{1}{\Upsilon} \log \left[ \check{F} \left( \exp \left[ \frac{1}{1+\rho} V_{1,t'}^{\text{ICT},G} - \kappa_{d1,t}^{k\text{ICT},hG} \right], \dots, \exp \left[ \frac{1}{1+\rho} V_{1,t'}^{\text{non-ICT},R} - \kappa_{dN,t}^{k\text{non-ICT},hR} \right] \right) \right]$ . The mean of this distribution is  $\frac{1}{\Upsilon} \log \check{F} + \Upsilon \gamma$ , where  $\gamma$  is Euler's constant. Thus the option value of moving – also known as the inclusive value of the entire choice set in the discrete choice literature – is given by:

$$\Phi_{d,t}^{k,h} = \frac{1}{\Upsilon} \log \left[ \sum_{o,s} \left( \sum_n \exp \left( \frac{1}{1+\zeta} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn} \right)^{\frac{\Upsilon}{v}} \right)^v \right], \quad (73)$$

where the term  $\Upsilon \gamma$  is constant over time and markets and hence can be normalized to 0. Substituting the

option value of moving into the individual worker migration problem in equation 25 yields the expected worker value in equation 27.

Mobility shares are given by the probability of choosing a specific market  $(o, s, n)$  to move to:

$$\begin{aligned}\mu_{do,t}^{ks,hn} &= \frac{\partial \Phi_{d,t}^{k,h}}{\partial \left( \frac{1}{1+\rho} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn} \right)} \\ &= \underbrace{\frac{\exp \left( \frac{1}{1+\zeta} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn} \right)^{\frac{\gamma}{v}}}{\sum_{n'} \exp \left( \frac{1}{1+\zeta} V_{o,t'}^{s,n'} - \kappa_{do,t}^{ks,hn'} \right)^{\frac{\gamma}{v}}}}_{\text{switching between production and research}} \cdot \underbrace{\frac{\left[ \sum_{n'} \exp \left( \frac{1}{1+\zeta} V_{o,t'}^{s,n'} - \kappa_{do,t}^{ks,hn'} \right)^{\frac{\gamma}{v}} \right]^v}{\sum_{o'} \sum_{s'} \left[ \sum_{n'} \exp \left( \frac{1}{1+\zeta} V_{o',t'}^{s',n'} - \kappa_{do',t}^{ks',hn'} \right)^{\frac{\gamma}{v}} \right]^v}}_{\text{migration across regions and sectors}}\end{aligned}$$

where the first equality comes from McFadden (1978). Note that given the correlation function, the individual worker mobility problem can also be recast as a nested discrete choice problem: each worker first chooses which region-sector to move to and then whether to supply their labor for research or production in that region-sector. In this formulation, we can interpret the second term as the probability of choosing region  $o$  and sector  $s$  among all region-sector alternatives and the first term as the conditional probability of choosing production or research given the region-sector choice. The conditional option value of choosing between production and research given the choice of region  $o$  and sector  $s$  is:

$$\Phi_{do,t}^{ks,h} = \frac{v}{\Upsilon} \log \left[ \sum_n \exp \left( \frac{1}{1+\zeta} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn} \right)^{\frac{\gamma}{v}} \right].$$

The unconditional option value across all regions, sectors, and occupations – as shown in equation 73 – is related to this conditional option value as follows:

$$\Phi_{d,t}^{k,h} = \frac{1}{v} \log \left[ \sum_{o,s} \exp \left( \Phi_{do,t}^{ks,h} \right)^v \right]. \quad (74)$$

Notice that when  $v = 1$  in the correlation function, draws of  $\epsilon$  are independent across markets, and the worker mobility shares and option value collapses to expressions of the form in Caliendo et al. (2019). By leveraging the correlation function and techniques from the foundational discrete choice literature, this proof *clarifies and generalizes dynamic migration in general equilibrium to a setting where preference shocks are correlated across markets*, i.e. drawn from a generalized extreme value distribution with any arbitrary correlation function  $\check{F}$  that satisfies the properties listed in McFadden (1978) and  $\mu$ -homogeneity in Ben-Akiva and Francois (1983).  $\square$

## C.8 Proposition 1 (Spatial and Sectoral Direction of Innovation)

The ratio of inventor real wages across regions in the same sector follows directly from the text. Additionally, the sectoral direction of innovation is governed by:

$$\frac{\omega_{r,t^*}^{k,R}}{\omega_{r,t^*}^{k',R}} = \underbrace{\frac{A_{t^*}^k}{A_{t^*}^{k'}}}_{\text{differences in fundamental research productivity across sectors}} \cdot \underbrace{\frac{T_{r,t^*}^k}{T_{r,t^*}^{k'}}}_{\text{relative technology levels}} \cdot \underbrace{\left( \frac{L_{r,t^*}^{k,G}}{L_{r,t^*}^{k',G}} \right)^\chi}_{\text{benefits of colocation of innovation and production}} \cdot \underbrace{\frac{\int_{t^*}^{\infty} e^{-\zeta(t-t^*)} \sum_{d=1}^N \frac{\phi_{rd,t^*}^k}{\lambda_{r,t^*}^k} \frac{X_{d,t}^k}{1+\theta} \frac{1}{P_{r,t}} dt}{\int_{t^*}^{\infty} e^{-\zeta(t-t^*)} \sum_{d=1}^N \frac{\phi_{rd,t^*}^{k'}}{\lambda_{r,t^*}^{k'}} \frac{X_{d,t}^{k'}}{1+\theta} \frac{1}{P_{r,t}} dt}}_{\text{expected market potential of an idea}}. \quad (75)$$

The first term captures the one-time increase in ICT research productivity at a national level (the first exogenous component of the ICT shock), the second term captures relative technology levels across sectors, the third term anchors the field of innovative activity to the sectoral composition of production in the region, and the fourth term captures differences in the expected market potential of ideas across sectors.

## C.9 Proposition 2 (Balanced Growth Path)

*Proof. Part (i):* On the balanced growth path, technology levels in all regions and sectors grow at the same rate  $g = \frac{\dot{T}_r^k(t)}{T_r^k}$   $\forall r, k$ , the exogenous bilateral diffusion lags  $\delta_{ro}$  are constant over time and the innovation rate is given by:

$$\lambda_r^k(t) = \gamma_r^k T_r^k(t) \implies \dot{\lambda}_r^k(t) = \gamma_r^k \dot{T}_r^k(t). \quad (76)$$

From equation 9 we know that the technology level at each time  $t$  is given by:

$$T_{o,t}^k = \sum_{r=1}^N T_{ro,t}^k = \sum_{r=1}^N \int_{-\infty}^t \Omega_{ro}(t-t^*)^{1-\rho} \cdot \lambda_r^k(t^*) dt^*.$$

Taking the derivative w.r.t.  $t$  yields:

$$\dot{T}_{o,t}^k = \sum_{r=1}^N \int_{-\infty}^t \frac{d\Omega_{ro}(t-t^*)^{1-\rho}}{dt} \cdot \lambda_r^k(t^*) dt^* + \Omega_{ro}(0)^{1-\rho} \lambda_r^k(t).$$

Now using integration by parts with  $u = \lambda_r^k(t^*)$  and  $dv = \frac{d\Omega_{ro}(t-t^*)^{1-\rho}}{dt}$  yields:

$$\begin{aligned} \dot{T}_{o,t}^k &= \sum_{r=1}^N \left[ -\lambda_r^k(t^*) \Omega_{ro}(t-t^*)^{1-\rho} \right]_{t^*=-\infty}^{t^*=t} + \int_{-\infty}^t \Omega_{ro}(t-t^*)^{1-\rho} \dot{\lambda}_r^k(t^*) dt^* + \Omega_{ro}(0)^{1-\rho} \lambda_l(t) \\ &= \sum_{r=1}^N \lim_{t^* \rightarrow -\infty} \lambda_r^k(t^*) \Omega_{ro}(t-t^*)^{1-\rho} + \int_{-\infty}^t \Omega_{ro}(t-t^*)^{1-\rho} \gamma_r^k \dot{T}_r^k(t^*) dt^* \\ &= \sum_{r=1}^N \gamma_r^k \int_{-\infty}^t \Omega_{ro}(t-t^*)^{1-\rho} \dot{T}_r^k(t^*) dt^*. \end{aligned}$$

Thus we have that:

$$\begin{aligned}
\dot{T}_{o,t}^k &= \sum_{r=1}^N \gamma_r^k \int_{-\infty}^t \Omega_{ro}(t-t^*)^{1-\rho} \frac{\dot{T}_r^k(t^*)}{T_r^k(t^*)} T_r^k(t^*) dt^* \\
&= g \sum_{r=1}^N \gamma_r^k \int_{-\infty}^t \Omega_{ro}(t-t^*)^{1-\rho} T_r^k(t) e^{-g(t-t^*)} dt^* \\
&= \sum_r \gamma_r^k T_{r,t}^k \int_{-\infty}^t g e^{-g(t-t^*)} \left[ 1 - e^{-\delta_{ro}(t-t^*)} \right]^{1-\rho} dt^* \\
&= \sum_r \gamma_r^k T_{r,t}^k \int_0^\infty g e^{-ga} \left[ 1 - e^{-\delta_{ro}(a)} \right]^{1-\rho} da
\end{aligned} \tag{77}$$

where  $e^{-\delta_{rr}(t-t^*)} \equiv 0$  and the last equality follows from a change of variable  $a = t - t^*$ . Note that  $\int_0^\infty g e^{-ga} \left[ 1 - e^{-\delta_{ro}(a)} \right]^{1-\rho} da$  is a constant.

Now building on Eaton and Kortum (2024), in matrix form we have:

$$g\mathbf{T}^k = \Delta^k(g)\mathbf{T}^k \tag{78}$$

where  $\mathbf{T}^k$  is an  $N \times 1$  vector with representative element  $T_r^k$  and  $\Delta^k(g)$  is an  $N \times N$  matrix with representative element:

$$\Delta_{ro}^k(g) = \gamma_r^k \int_0^\infty g e^{-ga} \left[ 1 - e^{-\delta_{ro}(a)} \right]^{1-\rho} da.$$

The aggregate growth rate is the Perron-Frobenius root of equation 78 with relative technology levels  $\mathbf{T}$  corresponding to the Perron-Frobenius eigenvector that is defined up to a scalar multiple. Thus, any arbitrary set of exogenous diffusion speeds  $\delta_{ro}$  and endogenous innovation rates  $\gamma_r^k$  delivers a balanced growth path with parallel growth at rate  $g$  with level differences in technology  $T$  across regions.

**Part (ii):** I now solve for the remaining variables in the economy on the BGP. Suppose that the distribution of workers across regions, sectors, and occupations are constant. Since the relative technology levels are also constant on the BGP, the trade shares given by equation 12 and the market condition given by equation 23 imply that relative wages across regions and sectors in production is constant. In particular, trade shares on the BGP are given by:

$$\pi_{od}^k = \sum_{r=1}^N \int_0^\infty \frac{\Omega_{ro}(a) (w_o^k \tau_{od}^k)^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N \Omega_{ro'}(a) (w_{o'}^k \tau_{o'd}^k)^{-\frac{\theta}{1-\rho}}} \cdot \frac{\gamma_r^k T_r^k e^{-g^k \cdot a} \left[ \sum_{o'=1}^N \Omega_{ro'}(a) (w_{o'}^k \tau_{o'd}^k)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}}{\sum_{r'=1}^N \gamma_{r'}^k T_{r'}^k \int_0^\infty \left[ \sum_{o'=1}^N \Omega_{r'o'}(a') (w_{o'}^k \tau_{o'd}^k)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} e^{-g^k \cdot a'} da'} da. \tag{79}$$

where  $T_r^k$  are the relative technology levels determined by equation 78.

On the BGP, the price index in each region and sector is given by:

$$P_{d,t}^k \overset{BGP}{=} \Gamma \left[ \sum_{r'=1}^N \gamma_{r'}^k T_{r',t}^k \int_0^\infty e^{-ga} \left[ \sum_{o'=1}^N \Omega_{r'o'}(a) \left( w_{o'}^k \tau_{o'd}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} da \right]^{-\frac{1}{\theta}} \quad (80)$$

Differentiating this expression w.r.t time, the growth rate of the sectoral price index is:

$$g_{P^k} = \frac{\dot{P}_{d,t}^k \overset{BGP}{=}}{P_{d,t}^k \overset{BGP}{=}} = -\frac{1}{\theta} \sum_{r'} \tilde{\pi}_{r'd} \frac{\dot{T}_{r',t}^k}{T_{r',t}^k} = -\frac{g^k}{\theta} \quad (81)$$

Since preferences are Cobb-Douglas across local sectoral final goods for each region, the aggregate price index is:

$$P_{d,t} = \check{\Gamma} \prod_k \left( P_{d,t}^k \right)^{\iota^k} \quad (82)$$

where  $\check{\Gamma}$  is a constant. The growth rate of the aggregate price index on the BGP is:

$$g_p = -\frac{1}{\theta} \sum_k \iota^k g^k \quad (83)$$

On the BGP, the expected value of an idea is:

$$\begin{aligned} \check{V}_{r,t^*}^k &= \frac{P_{r,t^*}}{1+\theta} \int e^{-\zeta(t-t^*)} \sum_d \frac{\phi_{rd,t^*t}}{\gamma_r^k T_{r,t^*}^k} \frac{X_{d,t}}{P_{r,t}} dt \\ &= \frac{\sum_d \phi_{rd} X_d}{1+\theta} \frac{1}{\gamma_r^k T_{r,t^*}^k} \int e^{-\zeta(t-t^*)} e^{-g_P(t-t^*)} dt \\ &= \frac{\sum_d \phi_{rd} X_d}{1+\theta} \frac{1}{\gamma_r^k T_{r,t^*}^k} \frac{1}{\zeta - g_P/\theta}. \end{aligned} \quad (84)$$

Thus  $\check{V}_{r,t^*}^k$  is falling at rate  $g^k$  on the BGP while inventor wages are constant and given by:

$$w_r^{k,R} = \frac{\sum_d \check{\pi}_{rd} X_d}{1+\theta} \frac{1}{L_r^{k,R}} \frac{1}{\zeta - g_P/\theta}. \quad (85)$$

I now solve for the growth rate of worker expected value. Let  $\exp(V_{d,t}^{k,h}) = \exp(\check{V}_d^{k,h}) e^{g_V t}$  and  $P_{d,t} = \tilde{P}_d e^{g_P t}$ , where  $\check{V}_d^{k,h}$  and  $\tilde{P}_d$  are the detrended value and price respectively. Thus given production worker

wages, inventor wages and local aggregate prices, worker expected value is given by:

$$\begin{aligned}
V_{d,t}^{k,h} &= \tilde{V}_d^{k,h} + g_v t \\
&= \log \left( \frac{w_d^{k,h}}{\tilde{P}_d} \right) - g_p t + \frac{1}{\Upsilon} \log \left[ \sum_o \sum_s \left( \sum_n \exp \left( \frac{1}{1+\zeta} [\tilde{V}_o^{s,n} + g_v t'] - \kappa_{do}^{ks,hn} \right)^{\frac{\Upsilon}{v}} \right)^v \right] \\
&= \log \left( \frac{w_d^{k,h}}{\tilde{P}_d} \right) - g_p t + \frac{1}{1+\zeta} g_v t' + \frac{1}{\Upsilon} \log \left[ \sum_o \sum_s \left( \sum_n \exp \left( \frac{1}{1+\zeta} \tilde{V}_o^{s,n} - \kappa_{do}^{ks,hn} \right)^{\frac{\Upsilon}{v}} \right)^v \right]
\end{aligned} \tag{86}$$

On the BGP, the growth rate must be the same on both sides of the equation. Thus, the growth rate of expected value is given by:

$$\begin{aligned}
g_v &= -g_p + \frac{1}{1+\zeta} g_v \\
\implies g_v &= -\frac{1+\zeta}{\zeta} g_p = \frac{1+\zeta}{\zeta} \frac{1}{\theta} \sum_k \iota^k g^k
\end{aligned} \tag{87}$$

where the first equality comes from  $\frac{1}{1+\zeta} g_v t' = \frac{1}{1+\zeta} g_v t + \frac{1}{1+\zeta} g_v$  in equation 86 since  $\mathbb{E}_t(t' - t) = 1$ , because the Poisson arrival rate of move possibilities is 1. The detrended expected value of workers is given by:

$$\tilde{V}_d^{k,h} = \log \left( \frac{w_d^{k,h}}{\tilde{P}_d} \right) + \frac{1}{1+\zeta} g_v + \frac{1}{\Upsilon} \log \left[ \sum_o \sum_s \left( \sum_n \exp \left( \frac{1}{1+\zeta} \tilde{V}_o^{s,n} - \kappa_{do}^{ks,hn} \right)^{\frac{\Upsilon}{v}} \right)^v \right]. \tag{88}$$

Substituting the decomposition of expected worker value in equation 86 into equation 28 in the main text, worker mobility shares are alternately given by:

$$\mu_{do}^{ks,hn} = \frac{\exp \left( \frac{1}{1+\zeta} \tilde{V}_o^{s,n} - \kappa_{do}^{ks,hn} \right)^{\frac{\Upsilon}{v}}}{\sum_{n'} \exp \left( \frac{1}{1+\zeta} \tilde{V}_o^{s,n'} - \kappa_{do,t}^{ks,hn'} \right)^{\frac{\Upsilon}{v}}} \cdot \frac{\left[ \sum_{n'} \exp \left( \frac{1}{1+\zeta} \tilde{V}_o^{s,n'} - \kappa_{do}^{ks,hn'} \right)^{\frac{\Upsilon}{v}} \right]^v}{\sum_{o'} \sum_{s'} \left[ \sum_{n'} \exp \left( \frac{1}{1+\zeta} \tilde{V}_{o'}^{s',n'} - \kappa_{do'}^{ks',hn'} \right)^{\frac{\Upsilon}{v}} \right]^v} \tag{89}$$

and are constant on the BGP. Hence, from equation 29, the distribution of workers across regions, sectors, and occupations are constant on the BGP.

Thus the balanced growth path of the economy is obtained, where workers, wages, migration shares, trade shares and innovation rates are constant, technology in each sector  $k$  and region is growing at rate  $g^k$  determined from equation 78, prices in each region is growing at rate  $g_p = -\frac{1}{\theta} \sum_k \iota^k g^k$  and expected worker value is growing at rate  $g_v = \frac{1+\zeta}{\zeta} \frac{1}{\theta} \sum_k \iota^k g^k$ .  $\square$

### Special Cases:

(i) When  $\theta = \sigma$  (such that  $\rho = 0$ ), we have the case of exponential diffusion where idea applicabilities do

not matter (**case NA** for no applicabilities). Equation (77) collapses to:

$$\begin{aligned}\dot{T}_{o,t}^k &= \sum_r \gamma_r^k T_{r,t}^k \int_0^\infty g e^{-ga} \left[ 1 - e^{-\delta_{ro}(a)} \right] da \\ &= \sum_r g \gamma_r^k T_{r,t}^k \left( \frac{1}{g} + \frac{1}{g + \delta_{ro}} \right) \\ &= \sum_r \left( \frac{\delta_{ro}}{g + \delta_{ro}} \right) \gamma_r^k T_{r,t}^k\end{aligned}$$

yielding equation (49) in Eaton and Kortum (2024). The trade shares and price index collapse to the canonical Eaton and Kortum (2002) expressions:

$$\pi_{od}^{NA} = \frac{T_{o,t} (w_o \tau_{od})^{-\theta}}{\sum_{o'=1}^N T_{o',t} (w_{o'} \tau_{o'd})^{-\theta}}, \quad P_{d,t}^{NA} = \Gamma \left[ \sum_{o'=1}^N T_{o',t} (w_{o'} \tau_{o'd})^{-\theta} \right]^{-\frac{1}{\theta}} \quad (90)$$

and idea diffusion shares are given by:

$$\varphi_{ro}^{NA} = \sum_{d=1}^N \frac{T_{ro,t} (w_o \tau_{od})^{-\theta}}{\sum_{o'=1}^N T_{o',t}^k (w_{o'} \tau_{o'd})^{-\theta}}. \quad (91)$$

All the other variables and growth rates remain the same as the full model.

(ii) When  $\Omega_{ro,t^*}(t - t^*) = \delta_{ro,t}$  for  $r \neq o$  and  $t \geq t^*$ , we have the case of instantaneous diffusion with idea applicabilities (**case ID**) as in Xiang (2023), equation (77) collapses to:

$$\begin{aligned}\dot{T}_{o,t}^k &= \sum_r \gamma_r^k T_{r,t}^k \int_0^\infty g e^{-ga} \delta_{ro}^{1-\rho} da \\ &= \sum_r \delta_{ro}^{1-\rho} \gamma_r^k T_{r,t}^k.\end{aligned}$$

The trade shares and price index collapse to the expressions in Ramondo and Rodríguez-Clare (2013):

$$\pi_{od}^{ID} = \sum_{r=1}^N \frac{(T_{ro,t})^{\frac{1}{1-\rho}} (w_o \tau_{od})^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N (T_{ro',t})^{\frac{1}{1-\rho}} (w_{o'} \tau_{o'd})^{-\frac{\theta}{1-\rho}}} \cdot \frac{\left[ \sum_{o'=1}^N (T_{ro',t})^{\frac{1}{1-\rho}} (w_{o'} \tau_{o'd})^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}}{\sum_{r'=1}^N \left[ \sum_{o'=1}^N (T_{ro',t})^{\frac{1}{1-\rho}} (w_{o'} \tau_{o'd})^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}} \quad (92)$$

$$P_{d,t}^{ID} = \Gamma \left[ \sum_{r'=1}^N \left[ \sum_{o'=1}^N (T_{r'o',t})^{\frac{1}{1-\rho}} (w_{o'} \tau_{o'd})^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \right]^{-\frac{1}{\theta}} \quad (93)$$

and the idea diffusion shares are:

$$\varphi_{od,t}^{ID} = \frac{(T_{ro,t})^{\frac{1}{1-\rho}} (w_o \tau_{od})^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N (T_{ro',t})^{\frac{1}{1-\rho}} (w_{o'} \tau_{o'd})^{-\frac{\theta}{1-\rho}}} \cdot \frac{\left[ \sum_{o'=1}^N (T_{ro',t})^{\frac{1}{1-\rho}} (w_{o'} \tau_{o'd})^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}}{\sum_{l'=1}^N \left[ \sum_{o'=1}^N (T_{ro',t})^{\frac{1}{1-\rho}} (w_{o'} \tau_{o'd})^{-\frac{\theta}{1-\rho}} \right]^{1-\rho}} \quad (94)$$

All the other variables and growth rates remain the same as the full model.

### C.10 Proposition 3 (Transition Path)

*Proof.* I first characterize the transition path with Assumption 3, then with Assumption 4 in order to apply dynamic hat algebra and map my model to data on innovation levels, trade shares, migration shares, and firm establishment networks.

#### (i) Transition Path with Continuous Innovation, Production, and Consumption; and Migration at Discrete Moments in Time (Assumption 3)

The transition path towards balanced growth is characterized by a trajectory of detrended expected values given by equation 88 and worker mobility shares given by equation 89 but with all variables varying over time, alongside the evolution of worker population given by equation 29:

$$\begin{aligned} \tilde{V}_{d,t}^{k,h} &= \int_t^{t'} \log \left( \frac{w_{d,\tilde{t}}^{k,h}}{\tilde{P}_{d,\tilde{t}}} \right) d\tilde{t} + \frac{1}{1+\zeta} g_v + \frac{1}{\Upsilon} \log \left[ \sum_o \sum_s \left( \sum_n \exp \left( \frac{1}{1+\zeta} \tilde{V}_{o,t'}^{s,n} - \kappa_{do}^{ks,hn} \right)^{\frac{\Upsilon}{v}} \right)^v \right] \\ \mu_{do,t}^{ks,hn} &= \frac{\exp \left( \frac{1}{1+\zeta} \tilde{V}_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn} \right)^{\frac{\Upsilon}{v}}}{\sum_{n'} \exp \left( \frac{1}{1+\zeta} \tilde{V}_{o,t'}^{s,n'} - \kappa_{do,t}^{ks,hn'} \right)^{\frac{\Upsilon}{v}}} \cdot \frac{\left[ \sum_{n'} \exp \left( \frac{1}{1+\zeta} \tilde{V}_{o,t'}^{s,n'} - \kappa_{do,t}^{ks,hn'} \right)^{\frac{\Upsilon}{v}} \right]^v}{\sum_{o'} \sum_{s'} \left[ \sum_{n'} \exp \left( \frac{1}{1+\zeta} \tilde{V}_{o',t'}^{s',n'} - \kappa_{do',t}^{ks',hn'} \right)^{\frac{\Upsilon}{v}} \right]^v} \\ L_{o,t'}^{s,n} &= \sum_h \sum_k \sum_d \mu_{do,t'}^{ks,hn} L_{d,t}^{k,h}, \end{aligned}$$

where at each time  $t$  the innovation and technology levels are given by equations 30 and 9 respectively:

$$\begin{aligned} \lambda_{r,t}^k &= A_{t^*}^k A_{r,t^*} \left( L_{r,t^*}^{k,G} \right)^\chi L_{r,t^*}^{k,R} T_{r,t^*}^k \\ T_{o,t}^k &= \sum_{r=1}^N \int_{-\infty}^t \Omega_{ro,t^*} (t-t^*)^{1-\rho} \cdot \lambda_{r,t^*}^k dt^*, \end{aligned}$$

the trade shares, idea adoption shares, price index, and market clearing condition are given by equations 12,

16, 18, and 23 respectively:

$$\begin{aligned}\pi_{od,t}^k &= \sum_{r=1}^N \int_{-\infty}^t \frac{\Omega_{ro,t^*}(t-t^*) \left(w_{o,t}^{k,G} \tau_{od,t}^k\right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'} \Omega_{ro',t^*}(t-t^*) \left(w_{o',t}^{k,G} \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}} \frac{\left[\sum_{o'} \Omega_{ro',t^*}(t-t^*) \left(w_{r,t}^{k,G} \tau_{ro',t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho} \lambda_{r,t^*}^k}{\sum_{r'} \int_{-\infty}^t \left[\sum_{o'} \Omega_{r'o',\check{t}}(t-\check{t}) \left(w_{o',t}^{k,G} \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho} \lambda_{r',\check{t}}^k d\check{t}} dt^* \\ \varphi_{ro,t}^k &= \sum_{d=1}^N \int_{-\infty}^t \frac{\Omega_{ro,t^*}(t-t^*) \left(w_{o,t}^{k,G} \tau_{od,t}^k\right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'} \Omega_{ro',t^*}(t-t^*) \left(w_{o',t}^{k,G} \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}} \frac{\left[\sum_{o'} \Omega_{ro',t^*}(t-t^*) \left(w_{r,t}^{k,G} \tau_{ro',t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho} \lambda_{r,t^*}^k}{\sum_{r'} \int_{-\infty}^t \left[\sum_{o'} \Omega_{r'o',\check{t}}(t-\check{t}) \left(w_{o',t}^{k,G} \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho} \lambda_{r',\check{t}}^k d\check{t}} dt^* \\ P_{d,t}^k &= \gamma \left[ \sum_{r'=1}^N \int_{-\infty}^t \left[ \sum_{o'=1}^N \Omega_{r'o',t^*}(t-t^*) \left(w_{o',t}^{k,G} \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r',t^*}^k dt^* \right]^{-\frac{1}{\theta}} \\ \frac{1+\theta}{\theta} w_{o,t}^{k,G} L_{o,t}^{k,G} &= \sum_d \pi_{od,t}^k \iota^k \left[ \sum_k \left( w_{d,t}^{k,G} L_{d,t}^{k,G} + \sum_r \varphi_{dr,t}^k \frac{1+\theta}{\theta} w_{r,t}^{k,G} L_{r,t}^{k,G} \right) \right],\end{aligned}$$

the wages of inventors are given by equation (22), the returns to innovation by equation (21), total expenditures by equation (43), and idea market shares by equation (44):

$$\begin{aligned}w_{r,t}^{k,R} &= \frac{\check{V}_{r,t}^k \lambda_{r,t}^k}{L_{r,t}^{k,R}} \\ \check{V}_{r,t}^k &= \int_t^\infty e^{-\zeta(t'-t)} \sum_{d=1}^N \frac{X_{d,t'}^k}{1+\theta} \cdot \frac{P_{r,t}}{P_{r,t'}} \cdot \frac{\check{\pi}_{ld}^k(t,t')}{\lambda_{l,t}^{k,\check{R}}} dt' \\ X_{d,t'}^k &= \iota^k \left[ \sum_k \left( w_{d,t'}^{k,G} L_{d,t'}^{k,G} + \sum_l \varphi_{dr,t'}^k \frac{1+\theta}{\theta} w_{r,t'}^{k,G} L_{r,t'}^{k,G} \right) \right] \\ \phi_{rd,tt'}^k &= \frac{\left[ \sum_{o'} \Omega_{lo',t}(t'-t) \left(w_{l',t'}^k \tau_{lo',t'}^k\right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{l,t}^{k,\check{R}}}{\sum_{l'} \int_{-\infty}^{t'} \left[ \sum_{o'} \Omega_{l'o',\check{t}}(t'-\check{t}) \left(w_{o',t'}^k \tau_{o'd,t'}^k\right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{l',\check{t}}^{k,\check{R}} d\check{t}}.\end{aligned}$$

## (ii) Transition Path with Innovation, Production, Consumption, and Migration Occuring Simultaneously at Discrete Moments in Time (Assumption 4)

I now apply Assumption 4 to characterize the transition path in time changes. Let  $\mathcal{T}$  be the set of times where innovation, production, consumption, and migration all occur simultaneously,  $\dot{x}_t = \frac{x_{t'}}{x_t}$  for any variable  $x$ , and  $u = \exp(V)$ . Then the detrended equations for the migration equilibrium (i.e. equations 88, 89, 29)

can be expressed in time changes:

$$\log \left( \widehat{u}_{d,t}^{k,h} \right) = \log \left( \frac{\widehat{w}_{d,t}^{k,h}}{\widehat{P}_{d,t}} \right) + \frac{1}{\Upsilon} \log \left[ \sum_n \left[ \sum_s \sum_o \mu_{do,t'}^{ks,hn} \left( \widehat{u}_{o,t'}^{s,n} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \widehat{\kappa}_{do,t}^{ks,hn} \right)^{\frac{\Upsilon}{v}} \right]^v \right] \quad (95)$$

$$\begin{aligned} \mu_{od,t'}^{ks,hn} &= \frac{\mu_{od,t}^{ks,hn} \left( \widehat{u}_{d,t'}^{s,n} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \widehat{\kappa}_{od,t}^{ks,hn} \right)^{\frac{\Upsilon}{v}}}{\sum_{n'} \mu_{od,t}^{ks,hn'} \left( \widehat{u}_{d,t'}^{s,n'} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \widehat{\kappa}_{od,t}^{ks,hn'} \right)^{\frac{\Upsilon}{v}}} \cdot \frac{\left[ \sum_{n'} \mu_{od,t'}^{ks,hn'} \left( \widehat{u}_{d,t'}^{s,n'} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \widehat{\kappa}_{od,t'}^{ks,hn'} \right)^{\frac{\Upsilon}{v}} \right]^v}{\sum_{n'} \left[ \sum_{d'} \sum_{s'} \mu_{od',t'}^{ks',hn'} \left( \widehat{u}_{d',t'}^{s',n'} \right)^{\frac{\Upsilon}{(1+\zeta)v}} \left( \widehat{\kappa}_{od',t'}^{ks',hn'} \right)^{\frac{\Upsilon}{v}} \right]^v} \end{aligned} \quad (96)$$

$$L_{d,t'}^{k,h} = \sum_n \sum_s \sum_o \mu_{od,t'}^{ks,hn} L_{o,t'}^{s,n}. \quad (97)$$

Technology levels is now given by:

$$T_{o,t}^k = \sum_{r=1}^N \sum_{t^* \in \mathcal{T}_{-\infty}} \Omega_{ro,t^*} (t - t^*)^{1-\rho} \cdot \lambda_{r,t^*}^k, \quad (98)$$

where the diffusion term  $\Omega_{ro,t^*}(t - t^*)$  remains unchanged, while innovation levels remain unchanged from equation (30). The trade shares, idea adoption shares, and price index is now given by:

$$\pi_{od,t}^k = \sum_{r=1}^N \sum_{t^* \in \mathcal{T}} \frac{\Omega_{ro,t^*} (t - t^*) \left( w_{o,t}^k \tau_{od,t}^k \right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'} \Omega_{ro',t^*} (t - t^*) \left( w_{o',t}^k \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}}} \frac{\left[ \sum_{o'} \Omega_{ro',t^*} (t - t^*) \left( w_{o',t}^k \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r,t^*}^k}{\sum_{r'} \sum_{t^* \in \mathcal{T}} \left[ \sum_{o'} \Omega_{r'o',\check{t}} (t - \check{t}) \left( w_{o',t}^k \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r',\check{t}}^k} \quad (99)$$

$$\varphi_{ro,t}^k = \sum_{d=1}^N \sum_{t^* \in \mathcal{T}} \frac{\Omega_{ro,t^*} (t - t^*) \left( w_{o,t}^k \tau_{od,t}^k \right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'} \Omega_{ro',t^*} (t - t^*) \left( w_{o',t}^k \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}}} \frac{\left[ \sum_{o'} \Omega_{ro',t^*} (t - t^*) \left( w_{o',t}^k \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r,t^*}^k}{\sum_{r'} \sum_{t^* \in \mathcal{T}} \left[ \sum_{o'} \Omega_{r'o',\check{t}} (t - \check{t}) \left( w_{o',t}^k \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r',\check{t}}^k} \quad (100)$$

$$P_{d,t}^k = \Gamma \left[ \sum_{r'=1}^N \sum_{t^* \in \mathcal{T}} \left[ \sum_{o'=1}^N \Omega_{r'o',t^*} (t - t^*) \left( w_{o',t}^k \tau_{o'd,t}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r',t^*}^k \right]^{-\frac{1}{\theta}}, \quad (101)$$

where  $\mathcal{T} \equiv \mathcal{T}_{-\infty}$  are the set of times from  $-\infty$  to  $t$  where innovation, production, consumption, and migration occurs. The market clearing condition remains unchanged from equation (23). Inventor wages are given by equation (22), total expenditures by equation (43), and the expected value of an idea and the idea market shares are now given by:

$$\check{V}_{r,t}^k = \sum_{t' \in \mathcal{T}_t^\infty} \left( \frac{1}{1+\zeta} \right)^{t'-t} \sum_{d=1}^N \frac{X_{d,t'}^k}{1+\theta} \cdot \frac{P_{r,t}}{P_{r,t'}} \cdot \frac{\phi_{rd,tt'}^k}{\lambda_{r,t}^k} \quad (102)$$

$$\phi_{rd,tt'}^k = \frac{\left[ \sum_{o'} \Omega_{ro',t} (t' - t) \left( w_{r',t}^k \tau_{ro',t'}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r,t}^k}{\sum_{r'} \sum_{\check{t} \in \mathcal{T}_{-\infty}^t} \left[ \sum_{o'} \Omega_{r'o',\check{t}} (t' - \check{t}) \left( w_{o',t}^k \tau_{o'd,t'}^k \right)^{-\frac{\theta}{1-\rho}} \right]^{1-\rho} \lambda_{r',\check{t}}^k}. \quad (103)$$

□

### Special Cases:

When idea applicabilities are not relevant (**case NA**), only changes in trade costs – as opposed to levels – along with the other fundamentals are required to simulate the transition path. This is because the trade shares, idea diffusion shares, and price index depend only on the contemporaneous technology stock rather than the trajectory of innovation in all past periods, and are given by equations 90 and 91:

$$\pi_{od,t}^{k,NA} = \frac{T_{o,t}^k \left( w_{o,t}^k T_{od,t}^k \right)^{-\theta}}{\sum_{o'=1}^N T_{o',t}^k \left( w_{o',t}^k T_{o'd,t}^k \right)^{-\theta}}, \quad \varphi_{ro,t}^{k,NA} = \sum_{d=1}^N \frac{T_{ro,t}^k \left( w_{o,t}^k T_{od,t}^k \right)^{-\theta}}{\sum_{o'=1}^N T_{o',t}^k \left( w_{o',t}^k T_{o'd,t}^k \right)^{-\theta}}, \quad P_{d,t}^{k,NA} = \Gamma \left[ \sum_{o'=1}^N T_{o',t}^k \left( w_{o',t}^k T_{o'd,t}^k \right)^{-\theta} \right]^{-\frac{1}{\theta}}.$$

Thus, the production side of the economy can also be expressed using dynamic hat algebra. Changes in trade shares, idea diffusion shares and the price index are given by:

$$\hat{\pi}_{od,t}^{k,NA} = \frac{\hat{T}_{o,t}^k \left( \hat{w}_{o,t}^k \hat{T}_{od,t}^k \right)^{-\theta}}{\sum_{o'=1}^N \hat{T}_{o',t}^k \left( \hat{w}_{o',t}^k \hat{T}_{o'd,t}^k \right)^{-\theta}} \quad (104)$$

$$\hat{\varphi}_{ro,t}^{k,NA} = \sum_{d=1}^N \frac{\hat{T}_{ro,t}^k \left( \hat{w}_{o,t}^k \hat{T}_{od,t}^k \right)^{-\theta}}{\sum_{o'=1}^N \hat{T}_{o',t}^k \left( \hat{w}_{o',t}^k \hat{T}_{o'd,t}^k \right)^{-\theta}} \quad (105)$$

$$\hat{P}_{d,t}^{k,NA} = \left[ \sum_{o'=1}^N \hat{T}_{o',t}^k \left( \hat{w}_{o',t}^k \hat{T}_{o'd,t}^k \right)^{-\theta} \right]^{-\frac{1}{\theta}} \quad (106)$$

and the market clearing condition remains unchanged from equation 23. Innovation levels remain unchanged from equation 30 and changes in technology levels are given by:

$$\begin{aligned} T_{o,t'}^k - T_{o,t}^k &= \sum_r \delta_{ro,t} \left( \Lambda_{l,t}^k - T_{lo,t}^k \right) \\ \implies \hat{T}_{o,t'}^k &= 1 - \delta_{ro,t} + \sum_l \frac{\delta_{lo,t} \Lambda_{l,t}^k}{T_{o,t}^k} \end{aligned} \quad (107)$$

where  $\Lambda_{r,t}^k = \sum_{t^* \in \mathcal{T}} \lambda_{r,t^*}^k$  is the stock of innovations produced in region  $r$  at time  $t$ . This recursive formulation of exponential idea diffusion comes from Eaton and Kortum (2024). Since changes in technology are a function of the previous technology level, data on initial technology levels are still required to solve the transition path quantitatively.

Similarly, when there is instantaneous diffusion (**case ID**), only changes in trade costs – as opposed to levels – along with the other fundamentals are required to simulate the transition path. Trade shares, idea

diffusion shares, and price index are given by equations 92 -94:

$$\begin{aligned}\pi_{od,t}^{k, ID} &= \sum_{r=1}^N \frac{\left(T_{ro,t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o,t}^k \tau_{od,t}^k\right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N\left(T_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o',t}^k \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}} \cdot \frac{\left[\sum_{o'=1}^N\left(T_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o',t}^k \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}}{\sum_{r'=1}^N\left[\sum_{o'=1}^N\left(T_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o',t}^k \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}} \\ P_{d,t}^{k, ID} &= \Gamma\left[\sum_{r'=1}^N\left[\sum_{o'=1}^N\left(T_{r'o',t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o',t}^k \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}\right]^{-\frac{1}{\theta}} \\ \varphi_{od,t}^{k, ID} &= \frac{\left(T_{ro,t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o,t}^k \tau_{od,t}^k\right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N\left(T_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o',t}^k \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}} \cdot \frac{\left[\sum_{o'=1}^N\left(T_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o',t}^k \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}}{\sum_{r'=1}^N\left[\sum_{o'=1}^N\left(T_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(w_{o',t}^k \tau_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}}\end{aligned}$$

Thus changes in trade shares, idea diffusion shares and the price index are given by:

$$\widehat{\pi}_{od,t}^{k, ID} = \sum_{r=1}^N \frac{\left(\widehat{T}_{ro,t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o,t}^k \widehat{\tau}_{od,t}^k\right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N\left(\widehat{T}_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o',t}^k \widehat{\tau}_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}} \cdot \frac{\left[\sum_{o'=1}^N\left(\widehat{T}_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o',t}^k \widehat{\tau}_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}}{\sum_{r'=1}^N\left[\sum_{o'=1}^N\left(\widehat{T}_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o',t}^k \widehat{\tau}_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}} \quad (108)$$

$$\widehat{P}_{d,t}^{k, ID} = \left[\sum_{r'=1}^N\left[\sum_{o'=1}^N\left(\widehat{T}_{r'o',t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o',t}^k \widehat{\tau}_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}\right]^{-\frac{1}{\theta}} \quad (109)$$

$$\widehat{\varphi}_{od,t}^{k, ID} = \frac{\left(\widehat{T}_{ro,t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o,t}^k \widehat{\tau}_{od,t}^k\right)^{-\frac{\theta}{1-\rho}}}{\sum_{o'=1}^N\left(\widehat{T}_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o',t}^k \widehat{\tau}_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}} \cdot \frac{\left[\sum_{o'=1}^N\left(\widehat{T}_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o',t}^k \widehat{\tau}_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}}{\sum_{r'=1}^N\left[\sum_{o'=1}^N\left(\widehat{T}_{ro',t}^k\right)^{\frac{1}{1-\rho}}\left(\widehat{w}_{o',t}^k \widehat{\tau}_{o'd,t}^k\right)^{-\frac{\theta}{1-\rho}}\right]^{1-\rho}} \quad (110)$$

and the market clearing condition remains unchanged from equation 23. Changes in technology levels are given by:

$$\begin{aligned}T_{o,t'}^k - T_{o,t}^k &= \sum_r \delta_{ro,t}^{1-\rho} \gamma_{r,t}^k T_{r,t}^k \\ \implies \widehat{T}_{o,t'}^k &= 1 + \sum_r \delta_{ro,t}^{1-\rho} \gamma_{r,t}^k \frac{T_{r,t}^k}{T_{o,t}^k}\end{aligned} \quad (111)$$

and innovation levels remain unchanged from equation 30.

### C.11 Proposition 4 (Regional and Aggregate Welfare)

*Proof.* I extend the welfare derivation and expression in Caliendo et al. (2019) to my setting, where: (i) preference shocks are correlated across markets; (ii) there is endogenous and microfounded innovation and technology diffusion, and (iii) the endogenous distribution of innovation and technology diffusion across markets drives parallel growth in all regions in the long run.

The expected worker value in equation (27) is given by:

$$V_{d,t}^{k,h} = U(C_{d,t}^{k,h}) + \frac{1}{\Upsilon} \log \left[ \sum_o \sum_s \left( \sum_n \exp \left( \frac{1}{1+\zeta} V_{o,t'}^{s,n} - \kappa_{do,t}^{ks,hn} \right)^{\frac{\Upsilon}{v}} \right)^v \right] = \log \left( \frac{w_{d,t}^{k,h}}{P_{d,t}} \right) + \Phi_{d,t}^{k,h}.$$

I now express the option value  $\Phi_{d,t}^{k,h}$  in terms of own-mobility shares:

$$\begin{aligned} \Phi_{d,t}^{k,h} &= -\frac{1}{\Upsilon} \log \mu_{dd,t}^{kk} + \frac{v}{\Upsilon} \log \left[ \sum_n \exp \left( \frac{1}{1+\zeta} V_{d,t'}^{k,n} - \kappa_{dd,t}^{kk,hn} \right)^{\frac{\Upsilon}{v}} \right] \\ &= -\frac{1}{\Upsilon} \log \mu_{dd,t}^{kk} - \frac{v}{\Upsilon} \log (\mu_{dd,t}^{kk,hh} | \mu_{dd,t}^{kk}) + \frac{v}{\Upsilon} \log \left[ \exp \left( \frac{1}{1+\zeta} V_{d,t'}^{k,h} \right)^{\frac{\Upsilon}{v}} \right] \\ &= \frac{1}{1+\zeta} V_{d,t'}^{k,h} - \frac{1}{\Upsilon} \log \mu_{dd,t}^{kk} - \frac{v}{\Upsilon} \log (\mu_{dd,t}^{kk,hh} | \mu_{dd,t}^{kk}). \end{aligned}$$

Hence, expected worker value can be alternately expressed as:

$$V_{d,t}^{k,h} = \log \left( \frac{w_{d,t}^{k,h}}{P_{d,t}} \right) + \frac{1}{1+\zeta} V_{d,t'}^{k,h} - \frac{1}{\Upsilon} \log \mu_{dd,t}^{kk} - \frac{v}{\Upsilon} \log (\mu_{dd,t}^{kk,hh} | \mu_{dd,t}^{kk}).$$

Iterating this equation forward, we have:

$$\begin{aligned} V_{d,t}^{k,h} &= \sum_{t' \in \mathcal{T}_t^\infty} \left( \frac{1}{1+\zeta} \right)^{t'-t} \log \left( \frac{w_{d,t}^{k,h}}{P_{d,t} \left( \mu_{dd,t}^{kk} \right)^{1/\Upsilon} \left( \mu_{dd,t}^{kk,hh} | \mu_{dd,t}^{kk} \right)^{v/\Upsilon}} \right) \\ &= \sum_{t' \in \mathcal{T}_t^\infty} \left( \frac{1}{1+\zeta} \right)^{t'-t} \left\{ \underbrace{\log \left( \frac{w_{d,t}^{k,h}}{\tilde{P}_{d,t}} \right)}_{\text{detrended future value}} - \underbrace{\log \left[ \left( \mu_{dd,t}^{kk} \right)^{1/\Upsilon} \left( \mu_{dd,t}^{kk,hh} | \mu_{dd,t}^{kk} \right)^{v/\Upsilon} \right]}_{\text{option value of movement}} \right\} + \underbrace{\frac{1+\zeta}{\zeta} \frac{1}{\theta} \sum_k \iota^k g^k}_{\text{long-run growth}}. \end{aligned}$$

Denote  $\hat{x}$  as the counterfactual of any variable  $x$ . My measure of the **welfare** impact in **market**  $(d, k, h)$  of an anticipated sequence of counterfactual changes in fundamentals from time  $t = 0$  is the compensating

variation in consumption for market  $(d, k, h)$  at  $t = 0$ ,  $\log \delta_d^{k,h}$  given by:

$$\acute{V}_{d,0}^{k,h} = V_{d,0}^{k,h} + \sum_{t' \in \mathcal{T}_0^\infty} \left( \frac{1}{1+\zeta} \right)^{t'} \log \delta_d^{k,h}.$$

Rearranging this equation and substituting the expressions for expected worker value, we have:

$$\begin{aligned} \log(\delta_d^{k,h}) &= \frac{\zeta}{1+\zeta} [\acute{V}_{d,0}^{k,h} - V_{d,0}^{k,h}] \\ &= \left(1 - \frac{1}{1+\zeta}\right) \sum_{t' \in \mathcal{T}_0^\infty} \left( \frac{1}{1+\zeta} \right)^{t'} \log \left( \frac{\acute{w}_{d,t}^{k,h} \tilde{P}_{d,t} \left( \mu_{dd,t}^{kk} \right)^{1/\gamma} \left( \mu_{dd,t}^{kk,hh} | \mu_{dd,t}^{kk} \right)^{v/\gamma}}{w_{d,t}^{k,h} \tilde{P}_{d,t} \left( \mu_{dd,t}^{kk} \right)^{1/\gamma} \left( \mu_{dd,t}^{kk,hh} | \mu_{dd,t}^{kk} \right)^{v/\gamma}} \right) + \frac{1}{\theta} \sum_k \iota^k (\acute{g}^k - g^k) \\ &= \sum_{t' \in \mathcal{T}_0^\infty} \left( \frac{1}{1+\zeta} \right)^{t'} \log \left( \underbrace{\frac{\widehat{w}_{d,t}^{k,h}}{\widehat{\tilde{P}_{d,t}}} \frac{1}{\left( \widehat{\mu_{dd,t}^{kk}} \right)^{1/\gamma} \left( \widehat{\mu_{dd,t}^{kk,hh}} | \widehat{\mu_{dd,t}^{kk}} \right)^{v/\gamma}}}_{\substack{\text{change in} \\ \text{future} \\ \text{detrended} \\ \text{real wages}}} \right) + \underbrace{\frac{1}{\theta} \sum_k \iota^k (\acute{g}^k - g^k)}_{\text{growth effects}}, \end{aligned}$$

where  $\widehat{x}_{t'} = \frac{\dot{x}_{t'}}{\dot{x}_t}$  denotes the counterfactual change in any variable  $x$ .

Using this measure, I define **local and aggregate welfare** as population-weighted averages of welfare in the relevant markets:

$$\begin{aligned} \log(\delta_d) &= \sum_{k,h} \frac{L_d^{k,h}}{\sum_{k,h} L_d^{k,h}} \log(\delta_d^{k,h}) \\ \log(\delta) &= \sum_{d,k,h} \frac{L_d^{k,h}}{\sum_{d,k,h} L_d^{k,h}} \log(\delta_d^{k,h}). \end{aligned}$$

□

## D Model Extensions

My quantitative spatial growth model in the main text is deliberately parsimonious to capture the main drivers of the rising spatial concentration of innovation from the data. Nonetheless, the central feature of my spatial model is that I introduce endogenous innovation and integrate it with technology diffusion at the *idea* level, the fundamental unit of the Eaton-Kortum world. Thus, my model requires minimal assumptions and can flexibly incorporate other components in quantitative dynamic spatial models.

### D.1 Dynamic Worker Sorting by Skill

To incorporate college-educated  $H$  and non-college educated  $S$  workers, equation 27-29 can be duplicated for each worker type. Equilibrium migration shares would now capture worker sorting patterns in the data as opposed to aggregate bilateral migration flows across both worker types. In the production of goods, labor

is now a composite of college-educated and non-college educated workers. For instance, one could assume a CES aggregate between both types of labor:

$$L_{o,t}^k = \left[ \left( L_{o,t}^{H,k} \right)^{\frac{\varphi-1}{\varphi}} + \left( L_{o,t}^{S,k} \right)^{\frac{\varphi-1}{\varphi}} \right]^{\frac{\varphi}{\varphi-1}} \quad (112)$$

such that wages  $w_{o,t}^k$  in the market clearing condition is now an aggregate of wages of each worker type, with no other required changes. Note that I drop the superscript  $G$  for production workers in this section of the appendix for notational clarity, since there is no discussion on innovation workers.

## D.2 Multiple Factors of Production and Input-Output Loops

Input-output loops can be easily incorporated into the trade equilibrium at each  $t$  following Alvarez and Lucas (2007); Caliendo and Parro (2015). Instead of just using labor, production of each variety  $\nu$  in each region is now given by a two-tier Cobb-Douglas constant returns to scale technology:

$$Y_{o,t}^k(\nu) = z_{o,t}^k(\nu) \left[ (K_{o,t}^k)^\psi (L_{o,t}^k)^{1-\psi} \right]^\chi M_{o,t}^{1-\chi} \quad (113)$$

where  $z_{o,t}(\nu)$  is the productivity drawn from the multivariate Fréchet distribution given by equation 7,  $K_{o,t}$  is capital used in production, referring to commercial structures such as local buildings,  $L_{o,t}$  is labor, and  $M_{o,t}$  is intermediate inputs purchased from the final goods producer in the same region,  $\psi$  is the share of local structures in value added, and  $\chi$  is the share of value added. The unit cost of an input bundle is given by:

$$x_{o,t}^k = \tilde{\Gamma} \left[ (\check{r}_{o,t})^\psi \left( w_{o,t}^k \right)^{1-\psi} \right]^\chi P_{o,t}^{k(1-\chi)} \quad (114)$$

where  $\tilde{\Gamma}_o$  is a constant,  $\check{r}_{o,t}$  is the rental rate of capital from local capitalists, and  $P_{o,t}$  is also the price of the local industry aggregate of varieties. Replacing  $w_{o,t}$  with  $x_{o,t}$  in equation 12 yields the equilibrium trade shares.

Capital market clearing is given by:

$$\check{r}_{o,t} K_{o,t}^k = \frac{1-\psi}{\psi} w_{o,t}^k L_{o,t}^k \quad (115)$$

Since capital income is a constant multiple of production worker income, the combined capital and labor market clearing condition is still given by equation 23 in the main paper. With input-output loops of the form in equation 114 – where goods producers purchase the final good only in that sector<sup>41</sup> – intermediate good spending is a constant multiple of production worker and capital income:

$$X_{o,t}^I = \frac{1-\chi}{\chi} \left( w_{o,t}^k L_{o,t}^k + r_{o,t} K_{o,t}^k \right) = \frac{1-\chi}{\chi} \frac{1}{\psi} w_{o,t}^k L_{o,t}^k. \quad (116)$$

Thus, the combined capital and labor market clearing condition is still given by equation 23 in the main

---

<sup>41</sup>If goods producers purchase the final goods from all sectors, the market clearing condition will be slightly modified, as shown in Caliendo and Parro (2015), but remains highly tractable.

paper. All the other equations in the model also remain the same.

### D.3 Capital Accumulation

Capital accumulation can be added following Kleinman et al. (2023). Apart from workers and local immobile firms, we can introduce local immobile capitalists. In each region, local immobile capitalists build durable local structures  $K_{o,t}$  and rent to firms in different sectors at a nominal rate  $\check{r}_{o,t}$ . With their rental income, capitalists choose their consumption and investment to maximize intertemporal utility:

$$\check{V}_{o,t}^K = \mathbb{E}_t \sum_{s=0}^{\infty} \beta^{t+s} \frac{(C_{o,t+s}^K)^{1-1/\eta}}{1 - 1/\eta} \quad (117)$$

subject to their budget constraint:

$$r_{o,t} K_{o,t} = P_{o,t} (C_{o,t}^K + K_{o,t+1} - (1 - \delta_{o,t}) K_{o,t}) \quad (118)$$

where rental income can be used for consumption (first term), or saving to increase future capital (last two terms). Per period consumption expenditures  $P_{o,t} C_{o,t}^K$  are allocated via the same utility function as households.

The optimal consumption and saving decisions are given by:

$$C_{o,t}^K = \varsigma_{o,t} R_{o,t} K_{o,t} \quad (119)$$

$$K_{o,t+1} = (1 - \varsigma_{o,t}) R_{o,t} K_{o,t} \quad (120)$$

$$\varsigma_{o,t}^{-1} = 1 + \beta^\eta \left( \mathbb{E}_t \left[ R_{o,t+1}^{\frac{\eta-1}{\eta}} \varsigma_{o,t+1}^{-\frac{1}{\eta}} \right] \right)^\eta \quad (121)$$

where the consumption rate  $\varsigma_{o,t}$  is defined recursively and the gross return on capital is  $R_{o,t} \equiv 1 - \delta + r_{o,t}/P_{o,t}$ . The other equations in the model remain the same.

### D.4 Amenities with Congestion and Agglomeration in Production

Amenities, explicit congestion forces, and agglomeration in production can be introduced following Allen and Arkolakis (2014). The instantaneous utility function is now given by:

$$U(C_{o,t}, B_{o,t}) = \log(B_{o,t} C_{o,t}) \quad (122)$$

with  $B_{o,t} = \check{B}_{o,t} L_{o,t}^{-\xi}$  where  $\check{B}_{o,t}$  are fundamental amenities in region  $o$  at time  $t$  and  $\xi$  captures congestion forces. The goods production function in equation 113 now becomes:

$$Y_{o,t}^k(\nu) = z_{o,t}^k(\nu) (L_{o,t}^k)^{\check{\alpha}} \left[ (K_{o,t}^k)^\psi (L_{o,t}^{k,G})^{1-\psi} \right]^\chi M_{o,t}^{1-\chi} \quad (123)$$

where  $\check{\alpha}$  captures agglomeration in production. All the other equations in the model remain the same.

## E Details of Model Quantification

### E.1 Simulation Algorithm

Using Propositions 2 and 3, we can solve for the transition and balanced growth paths given an initial allocation of the economy and an anticipated convergent sequence of fundamentals. Since I use annual data on trade flows, migration flows, and innovation levels, the intervals between two innovation, production, or migration times are 1. Let superscripts in parentheses  $(i)$  denote guesses for variables I search over in iteration  $i$  through the loop. The numerical algorithm is as follows:

- **Step 1:** Guess a path of  $\{\hat{u}_{o,t+1}^{k,h,(i)}\}_{t=0}^T$  that converges to  $\hat{u}_{o,T+1}^{k,h,(i)} = 1$ .
- **Step 2:** For all  $t \geq 0$ , use equation (35), the guesses for  $\{\hat{u}_{o,t+1}^{k,h,(i)}\}_{t=0}^T$  and data on initial migration shares  $\mu_{od,-1}^{ks,ha}$  and the trajectory of exogenous changes in migration costs  $\{\hat{\kappa}_{od,t}^{ks,ha}\}_{t=0}^T$  to solve for the trajectory of migration shares  $\{\mu_{od,t}^{ks,ha}\}_{t=0}^T$ .
- **Step 3:** Use equation (36), the trajectory of migration shares  $\{\mu_{od,t}^{ks,ha}\}_{t=0}^T$  and the initial labor allocation  $\{L_{o,0}^{k,h}\}$  to solve for the trajectory of labor allocations  $\{L_{o,t}^{k,h}\}_{t=0}^T$ .
- **Step 4:** Use equations (30) and (37), the trajectory of inventors  $\{L_{o,t}^{k,\check{R}}\}_{t=0}^T$  and data on initial technology levels  $\{T_{o,0}^k\}$  and the trajectory of exogenous fundamental research productivities  $\{A_{o,t}^{k,\check{R}}\}$  to solve for the trajectory of innovation and technology levels  $\{\lambda_{o,t}^{k,\check{R}}, T_{o,t}^k\}_{t=0}^T$ .
- **Step 5:** Use equation (33), innovation levels in the final period  $\lambda_{o,T}^{k,\check{R}}$ , and data on the exogenous diffusion lag in the final period  $\delta_{lo,T}$  to solve for the growth rate of technology in each sector  $g^k$  and hence the growth rate of prices  $g_p = \frac{1}{\theta} \sum_k \iota^k g^k$  on the **balanced growth path**.
- **Step 6:** Use the trajectory of labor allocations and innovation levels  $\{L_{o,t}^{k,\check{R}}, \lambda_{o,t}^{k,\check{R}}\}_{t=0}^T$  from Steps 3 and 4 respectively and data on the trajectory of exogenous idea diffusion speeds and changes in trade costs  $\{\delta_{lo,t}, \hat{\tau}_{od,t}^k\}_{t=0}^T$  to solve the **trade equilibrium** at each  $t$ :
  - **Step 6A:** Guess a value for  $w_{o,t}^{k,(j)}$ , and use equations (38), (39) and (23) to get  $w_{o,t}^{k,(j+1)}$  and repeat until convergence
  - **Step 6B:** Use the converged value for wages  $w_{o,t}^k$  to compute the sectoral price index from equation (40), total expenditures in each sector  $X_{d,t}^k$  from equation (43) and idea market shares  $\check{\pi}_{ld}^k(t^*, t)$  from equation (13).
- **Step 7:** Use the trajectory of total expenditures and idea market shares  $\{X_{d,t}^k, \check{\pi}_{ld}^k(t^*, t)\}_{t=0, t^* \leq t}^T$  and

equation (42) to compute the trajectory of the value of ideas  $\{\check{V}_{o,t}^k\}_{t=0}^T$ , then equation (41) to compute the trajectory of inventor wages  $\{w_{o,t}^{k,\check{R}}\}_{t=0}^T$ .

- **Step 8:** For each  $t \geq 1$ , use equation (34),  $\{\hat{u}_{o,t+2}^{L(i)}\}$  from the previous outer loop iteration, migration shares  $\{\mu_{od,t}^{ks,ha}\}_{t=0}^T$  from Step 2, production worker wages and sectoral price indices from Step 6, inventor wages  $\{w_{o,t}^{k,\check{R}}\}_{t=0}^T$  from Step 7, and the growth rate of prices  $g_p$  from Step 4 to solve backwards in time for  $\{\hat{u}_{o,t+1}^{k(i+1)}\}_{t=0}^T$ .
- **Step 9:** If  $\{\hat{u}_{o,t+1}^{k,h,(i)}\}_{t=0}^T \approx \{\hat{u}_{o,t+1}^{k,h,(i+1)}\}_{t=0}^T$ , i.e. the maximum difference across all  $t$  is less than some pre-specified tolerance level,  $\{\hat{u}_{o,t+1}^{k,h,(i)}, \mu_{od,t}^{ks,ha}, L_{o,t}^{k,h}\}_{t=0}^T$  is the solution to the problem. Otherwise update the initial guess and repeat the steps until convergence.