Technological Waves, Knowledge Diffusion, and Local Growth*

Enrico Berkes† Ruben Gaetani‡ Martí Mestieri§

5th September 2024

Abstract

We develop a spatial model of endogenous growth with frictional knowledge diffusion to examine the effect of technological waves—defined as long-term shifts in the importance of specific knowledge fields—on local growth dynamics. We calibrate the model using a new dataset of historical geolocated patents spanning over one hundred years. We find that frictions to idea diffusion across locations and technological fields account for two-thirds of the empirical relationship between exposure to technological waves and local growth in the United States during the twentieth century. Counterfactual experiments suggest that future changes in the technological landscape may have substantial geographical effects.

Keywords: Cities, Population Growth, Technology Diffusion, Innovation, Patents.
JEL Classification: R12, O10, O30, O33, O47.

*This paper previously circulated as "Cities and Technological Waves" (CEPR DP16794, 2021) and "Technological Waves and Local Growth". We thank the Editor, three anonymous referees, Mike Andrews, Paco Buera, Klaus Desmet, Ed Glaeser, Xian Jiang, Ben Jones, Jeffrey Lin, Sara Mitchell, Joel Mokyr, Nicola Persico, Frédéric Robert-Nicoud, Bruce Weinberg, Ben Zou, and participants at seminar and conference presentations for helpful comments and discussions. Gaetani gratefully acknowledges financial support from the Institute for Management and Innovation Research Grant. All errors are our own.

†University of Maryland, Baltimore County, enrico.berkes@umbc.edu.
‡University of Toronto, ruben.gaetani@utoronto.ca.
§IAE-CSIC, BSE, and CEPR, mestieri.marti@gmail.com.

1 Introduction

Recent theories of economic growth highlight the role of learning and knowledge diffusion as key productivity drivers (Buera and Lucas, 2018). Idea diffusion is highly localized, making geographical proximity a key determinant of people's capacity to learn and adopt existing knowledge, as widely demonstrated in empirical research (e.g., Jaffe et al., 1993; Greenstone et al., 2010). This paper investigates the role played by this diffusion process in shaping heterogeneous local growth dynamics across cities and regions.

Changes in the technological environment have historically coincided with periods of transformation of the economic geography. In the United States, this is exemplified by the rise and fall of industrial cities in the Rust Belt in the twentieth century and the emergence of innovation hubs specializing in information technology (IT) and pharmaceuticals in recent years (Glaeser and Gottlieb, 2009; Moretti, 2012). While these large fluctuations are linked to a variety of outcomes relevant to welfare, the economic channels that connect changes in the technological landscape to these patterns remain largely unexplored. Progress on this issue has been hampered by a lack of comprehensive data on the geography of innovation over long periods. Additionally, identifying specific channels has proven difficult both empirically and theoretically, due to challenges in using quasi-experimental approaches to study long-term outcomes and the inherent intractability of knowledge diffusion in spatial models.

This study presents new facts and theory linking changes in the technological environment and local growth dynamics. Using a novel dataset of geolocated historical U.S. patents spanning more than 100 years, we document that cities that are more favorably exposed to changes in the technological landscape—referred to as "technological waves"—systematically experience higher growth over the long-term. Motivated by this finding, we develop a model that integrates a spatial equilibrium setting into a theory of economic growth with frictional idea diffusion. The model formalizes the relationship between technological waves and local growth, allowing us to estimate the contributions of the underlying economic forces. The quantitative results suggest that frictions to idea diffusion explain nearly two-thirds of the empirical relationship between exposure to technological waves and local growth throughout the last century.

To begin the analysis, we use our historical data to show that the prominence of various areas of knowledge, measured as the share of patents from each field in the overall innovation landscape, changes slowly but significantly over time. We refer to these long-term changes as "technological waves." Various factors, including scientific advancements, demand changes, or new regulations can cause these shocks. In the paper, we do not take a stand on the origin of these shocks, and focus instead on their effects on local growth dynamics.

Using a measure of local exposure to technological waves, we document that cities with innova-

tion activities initially concentrated in expanding fields experience systematically higher population growth than cities with innovation activities concentrated in declining fields.[1] We also document that knowledge diffusion, measured by patent citations, is persistently localized, both geographically and technologically. These facts suggest that a city's ability to seize new technological opportunities is contingent on the pre-existing local availability of complementary ideas, and can explain why changes in the technological environment result in the success of some cities and the decline of others.

We formalize this mechanism in a spatial model of endogenous growth with innovation and frictional idea diffusion, which we use to quantify our proposed channel. In the model, agents make migration and occupational decisions based on their expected lifetime productivity in each location and sector. Productivity is determined by an individual choice between imitating an idea from the local knowledge distribution or innovating by enhancing an existing idea from a distribution comprising all locations and sectors. The applicability of this idea is affected by diffusion frictions and technological wave shocks, which we model as exogenous sectoral shifters that change the efficiency with which ideas from specific fields can be used as inputs for innovation. The existence of frictions to diffusion implies that ideas originating from geographically and technologically closer sources can be converted into new inventions more efficiently, making the local growth trajectory sensitive to technological wave shocks in a manner that favors locations with a greater availability of complementary ideas.

The framework remains tractable for any number of locations, sectors, and time periods and has a unique equilibrium with an explicit solution. Absent technological wave shocks, the model features a unique balanced growth path (BGP). The distribution of ideas for each location-sector endogenously retains a Fréchet structure, with an intuitive law of motion for its average productivity.

To illustrate the model's mechanics, we log-linearize the equilibrium conditions around the BGP and characterize the relationship between technological waves and the evolution of local productivity and population. We show that a measure of local exposure to technological waves relative to the overall economy is a sufficient statistic for predicting local population growth. In the special case of knowledge flows across sectors being of second-order importance relative to flows within sectors, this measure of exposure becomes a standard shift-share variable, providing

[1] As is common in the literature on long-term spatial growth (e.g., Voigtländer and Voth, 2013; Hanlon and Heblich, 2022), our outcome of interest is changes in local population, which we observe consistently throughout our sample. The literature has found that fluctuations in local population carry major implications for welfare. Periods of rapid population decline are associated with deteriorating local economic conditions and lead to lower tax revenues, lower capacity to provide public services, and a drop in various measures of economic well-being (Glaeser and Gyourko, 2005; Owens III et al., 2020). Periods of rapid population growth, while reflecting thriving economic conditions for workers and firms, also carry ambiguous welfare implications, as they can lead to increasing congestion, housing shortages, and higher economic segregation (Baum-Snow, 2023; Berkes and Gaetani, 2023).

a simple micro-foundation for shift-share empirical strategies widely used in the literature.

For the quantitative analysis, we extend the model to include empirically relevant dimensions such as overlapping generations, migration costs, and endogenous congestion forces, and compute its full non-linear dynamics. We set up the calibration so that the model exactly matches data on population and patenting between 1890 and 2010. We quantify frictions to knowledge diffusion by deriving a gravity equation for idea flows and estimating it using patent citations data, and calibrate technological waves by recovering the full path of exogenous shocks that rationalize patenting in each sector throughout the sample. We show that the calibrated model successfully captures key moments of interest, even if not directly targeted, such as the relationship between city size and average city income, and the variation in average income across locations and sectors.

In our main counterfactual, we compare the dynamics of the economy under the calibrated path of technological waves (which replicates the data by construction) with the dynamics that would have occurred if technological waves had stayed constant at their values in the initial period of the sample. Our results suggest that technological waves account for a substantial portion of the variation in city-level population growth over the past century. The endogenous mechanism of frictional knowledge diffusion accounts for almost two-thirds of the reduced-form relationship between technological waves and local growth. Frictions to knowledge diffusion across geographical areas and technological fields are both significant, with each of them accounting for roughly half of the overall effect. Technological waves contributed to delineate the current economic geography of the United States, penalizing growth in cities in the Central States, and fostering the growth of modern innovation hubs.

Our mechanism also implies that cities with greater specialization experience significantly more volatile local growth in response to technological wave shocks. We quantify this effect through Monte Carlo simulations of counterfactual paths of sectoral shocks, and find a sizable effect of specialization on volatility. Increasing specialization from the level of Boston (a highly diversified city) to the level of Austin (a highly specialized one) increases the standard deviation of local growth over 40 years by 4.59 percentage points.

Finally, we use the quantitative model to explore the predicted local growth dynamics in the coming decades under various technological scenarios. First, we project population growth across cities following a significant drop in frictions to diffusion across locations (due to the increasing availability of tools of remote collaboration) and across technological fields (due to advances in Artificial Intelligence improving possibilities of recombination of ideas from different knowledge areas). Second, we study the geographical effects of some plausible scenarios of technological trends affecting specific knowledge areas, including transportation, pharmaceuticals, and agriculture. We uncover large and heterogeneous geographical effects of these scenarios, which in some cases result in a reversal of fortune for some of the currently most successful innovation hubs.

Related Literature This paper draws on multiple strands of previous research. First, our theory is based on modeling idea flows across locations and sectors, with frictions in knowledge diffusion playing an important role in explaining city dynamics. While a large body of literature has documented the strength and geographical reach of localized knowledge spillovers (e.g., Jaffe et al., 1993; Audretsch and Feldman, 1996; Greenstone et al., 2010), their role in long-run city dynamics is still largely unexplored.

The complexity of modeling idea diffusion in a spatial setting has long been a primary obstacle to this assessment. However, in recent years, two bodies of literature have provided tools making this problem more tractable. First, many papers have developed endogenous growth models that highlight knowledge recombination, imitation, and diffusion (e.g., Lucas and Moll, 2014; Perla and Tonetti, 2014; De la Croix et al., 2018; Buera and Oberfield, 2020; Huang and Zenou, 2020; Lind and Ramondo, 2023). Second, the literature on quantitative spatial economics has developed tools for analyzing the spatial distribution of economic activity, both within (e.g., Ahlfeldt et al., 2015; Heblich et al., 2020) and across cities (e.g., Allen and Arkolakis, 2014; Desmet et al., 2018b; Fajgelbaum and Gaubert, 2020).[2] This paper combines insights from these two literatures and develops a tractable endogenous growth model in a spatial economy that can be quantitatively disciplined using long-term data on city population and patents.

Several existing papers have studied innovation and knowledge flows in firm and industry dynamics (e.g., Kogan et al., 2017; Akcigit and Kerr, 2018; Cai and Li, 2019; Atkeson and Burstein, 2019) and have developed static models that emphasize localized knowledge spillovers as determinants of the economic geography (e.g., Davis and Dingel, 2019). This study connects these areas of the literature by quantitatively assessing the importance of frictions to idea diffusion for city dynamics. Our paper is also related to recent work on the relationship between spatial and aggregate growth. The paper in this literature which is closest to ours is Cai et al. (2022b), which examines spatial growth in a model with idea diffusion through trade and migration. In our paper, we maintain a reduced-form view of what constitutes a channel of diffusion (while introducing a sufficiently flexible parametrization to encompass a wide range of possible channels), and focus instead on how the underlying frictions shape the local response to technological waves.[3]

There is a vast literature studying the long-run evolution of the economic geography, with a focus on path dependence and reversals of fortune (e.g., Brezis and Krugman, 1997; Davis and Weinstein, 2002; Bleakley and Lin, 2012; Kline and Moretti, 2014; Allen and Donaldson, 2022), and on the effects of aggregate, regional, or sectoral shocks (e.g., Desmet et al., 2018a; Caliendo et al., 2018). Recent studies have emphasized channels such as migration frictions (Hornbeck

[2] Buera and Lucas (2018) and Redding and Rossi-Hansberg (2017) provide comprehensive reviews of the literatures on models of endogenous growth with idea flows and on quantitative spatial equilibrium models, respectively.

[3] Other contributions to this literature include Arkolakis et al. (2020) and Burchardi et al. (2020), who develop models of spatial growth to quantify the contribution of international migration to aggregate growth.

and Moretti, 2018; Borusyak et al., 2022), commuting across locations (Monte et al., 2018), and general equilibrium effects (Adao et al., 2020). In our paper, we focus on shocks that, regardless of their origin, affect innovation possibilities and returns in different technological areas, and we isolate the role that frictions to knowledge diffusion play in shaping the spatial response to those shocks, while allowing other residual channels to contribute to this response. While the focus on innovation and frictional idea diffusion is new to this literature, a large body of research has analyzed the historical dynamics of U.S. geography from both an empirical (e.g., Bostic et al., 1997; Simon and Nardinelli, 2002; Michaels et al., 2012; Desmet and Rappaport, 2017) and a structural and quantitative perspective (e.g., Duranton, 2007; Desmet and Rossi-Hansberg, 2014; Nagy, 2017; Eckert and Peters, 2022; Morris-Levenson and Prato, 2021; Kleinman et al., 2021; Giannone, 2022).

This paper also contributes to the ongoing debate regarding the returns to local specialization (Marshall, 1890) and diversity (Jacobs, 1969). Contributions to this literature include Glaeser et al. (1992), who find evidence supporting Jane Jacob's view of urban diversity as a key driver of local employment growth, and Duranton and Puga (2001), who develop a model in which diversified and specialized cities coexist in equilibrium.[4] This paper proposes and quantifies a new mechanism by which diversification affects long-term city growth by influencing a city's responsiveness to changes in the technological landscape.[5] By doing so, the model provides a new lens for interpreting the effect of local policies designed to increase local diversity.

The remainder of the paper is organized as follows. Section 2 introduces the data and presents historical trends and the motivational facts on the relationship between city growth and the technological landscape. Section 3 presents the stylized model used to derive the theoretical predictions. Section 4 describes the model's extensions for quantitative analysis and the calibration. Section 5 presents the quantitative results. Section 6 concludes by discussing avenues for future research and the policy implications of our results.

2 Data and stylized facts

Our unit of analysis is the 1990 commuting zone (CZ), which we keep fixed over time.[6] Throughout the paper, we refer to cities and CZs interchangeably.

[4] Holmes and Stevens (2004) provide an overview of the patterns of specialization in the United States.

[5] Consistently with this interpretation, Balland et al. (2015) find that cities with more diverse knowledge bases are less sensitive to technological crises, defined as sustained declines in patenting activity.

[6] Assuming a stable geography allows us to abstract from annexations and redefinition of town borders.

2.1 Data

We collect patent data from the Comprehensive Universe of U.S. Patents, or CUSP, to quantify innovative activities at the city level (Berkes, 2018). The CUSP contains data on the vast majority of patents issued by the U.S. Patent and Trademark Office between 1836 and 2015, with an estimated coverage of more than 90%. From the CUSP, we gather information regarding technology classes, filing date, and location of each inventor listed on a patent.[7] This is the first study to exploit the entire geolocated patent time series provided by the CUSP. Patents with multiple inventors are assigned fractionally to each listed CZ.[8]

We also collect data on population, human capital, and industry composition at the CZ level using decennial censuses for each decade between 1870 and 2010 from the Integrated Public Use Microdata Series (IPUMS, Ruggles et al., 2021) and the National Historical Geographic Information System (NHGIS, Manson et al., 2021). We build a consistent measure of the local density of human capital that combines available information on literacy and education. To make this variable comparable across decades, we construct an index based on the ranking of the relevant measure in each decade. Appendix B provides additional information regarding the data construction.

Our unit of time corresponds to 20-year intervals from 1870 to 2010. Patent counts are obtained by adding patents filed in the two decades surrounding the focal year (for instance, patents in the 1990 observation correspond to the total count between 1980 and 1999). We limit our sample to the subset of CZs in the contiguous United States that accounted for at least 0.01% of the total population for each decade since 1890.[9] This provides a sample of 485 CZs, which accounted for approximately 94.2% of the U.S. population in 2010. Sectors are defined as the class-groups obtained by clustering 3-digit International Patent Classification (IPC) categories into 11 class-groups, as detailed in Appendix Table A.1.[10]

2.2 Historical trends

During the last 150 years, the technological landscape has undergone significant transformations, as measured by changes in the prominence of various patent classes. These changes are already

[7]The CUSP assigns patents to the inventors' city of residence, regardless of the county listed in the patent's text. This allows us to build geographically consistent measures of innovation at the level of CZs for the extensive time period covered by our study.

[8]Berkes (2018) provides details about data collection, as well as summary statistics and stylized facts related to the underlying data. Andrews (2021), in a comparison of historical patent data, describes it as "currently the gold standard both in terms of completeness and scope of the types of patent information it contains" (p. 24).

[9]This rule requires that commuting zones had a population of at least 5,780 people in 1890 and 31,789 people in 2010. We further drop two commuting zones (with a combined population of less that 100,000 in 1990) which, having a patent count of zero in at least one of the 20-year periods in the sample, cannot be consistently included in the model calibration. None of the results is sensitive to alternative ways of dealing with these two observations.

[10]Patents listing multiple three-digit IPC classes are assigned fractionally to class-groups, in proportion to the frequency of each class-group in the list of classes.

evident when comparing patenting output across the broadest IPC classes (which are coarser than our baseline 11 categories). The bottom-right panel of Figure 1 illustrates the evolution of patenting shares since 1870.[11] The evolution of patenting shares over time is remarkably slow, highlighting the importance of using data that span a long period. The proportion of patents in the "Human Necessities" category, which includes innovations in agriculture and medical sciences, decreased in the early twentieth century, as agriculture lost its centrality to classes associated with the heavy manufacturing industry, such as "Transporting" and "Mechanical Engineering." In recent decades, "Human Necessities" patents have increased as innovation in medicine has gained prominence. In the second half of the twentieth century, patents in "Physics" and "Electricity" grew in importance, comprising more than 50% of the total innovation output in 2010.[12]

Not only has the composition of patenting changed significantly over time, but it also varies significantly between cities at any given time. Log-patenting shares residualized with respect to decade-class fixed effects have a standard deviation of 0.66 over 1890-2010. This large variation is illustrated in Appendix Figure A.2, which shows the distribution of patenting shares across cities in 1890 and 2010 (the beginning and end of our quantitative analysis, respectively) for each of the main IPC classes.

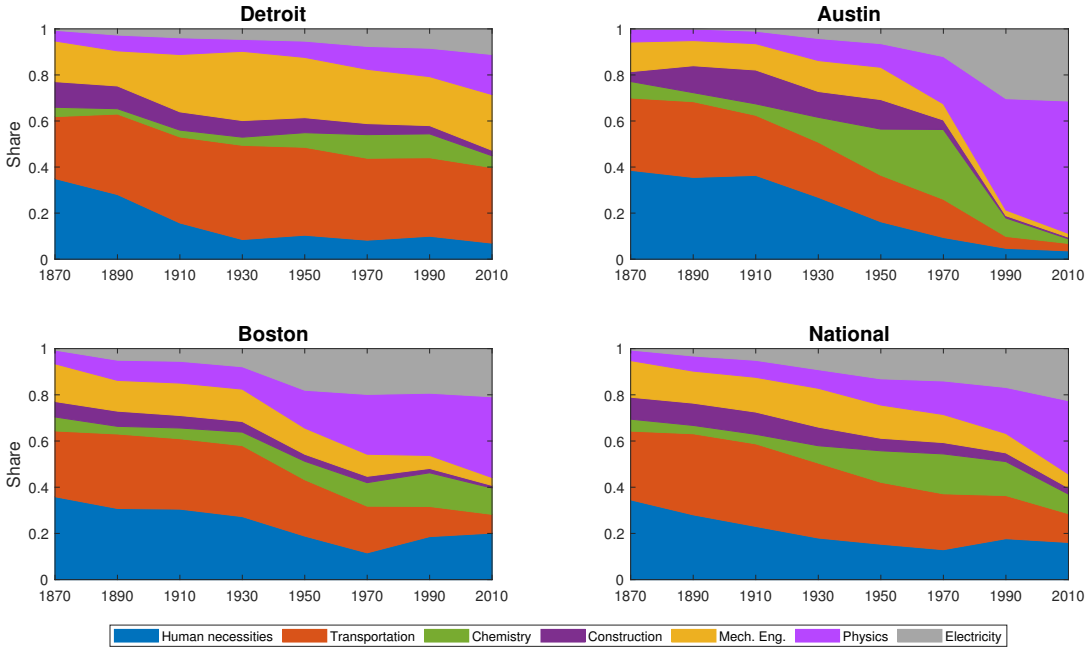
The remaining panels of Figure 1 depict three archetypal examples of this heterogeneity. Since the early 1900s, Detroit (top-left) has specialized in the production of patents related to "Transporting" and "Mechanical Engineering". In 1910, these two categories accounted for approximately 43% of its patent portfolio. Since the 1990s, there has been a small shift toward patents in "Physics" and "Electricity," but this pattern has remained largely unchanged throughout the century. Austin's (top-right) innovation activities were fairly diverse until the 1970s, when the proportion of patents in the classes "Physics" and "Electricity" began to increase, reaching 89% of the portfolio by 2010. By contrast, Boston (bottom-left) exhibits a diversified patenting output that has closely tracked the national trends throughout the decades.

In this paper, we argue that the heterogeneity in the composition of local knowledge makes cities differentially suited to capitalize on new innovation opportunities. The central premise is that knowledge is predominantly localized and diffuses slowly. This makes the trajectories of cities sensitive to changes in the technological landscape, as cities' current knowledge portfolio determines the extent to which they can exploit new technological opportunities. Consequently,

[11] For expositional purposes, we report results for the seven main IPC classes (that correspond to the first letter in the IPC). This has the drawback of bundling together, among others, innovations related to agriculture and medicine. Appendix Figure A.1 shows the corresponding distribution across the 11 IPC class-groups described in Appendix Table A.1 that we use in our analysis, which separates, among others, agriculture and medicine. Class names are abbreviated for clarity. The full description of each class can be found at https://www.wipo.int/classifications/ipc/en/. Kelly et al. (2021) provide an alternative measure of technological importance by constructing technology indices based on textual analysis of patent data.

[12] Classes "Physics" and "Electricity" include the bulk of innovation related to computers, electronics, and information and communication technology.

Figure 1: **Composition of the technological output**



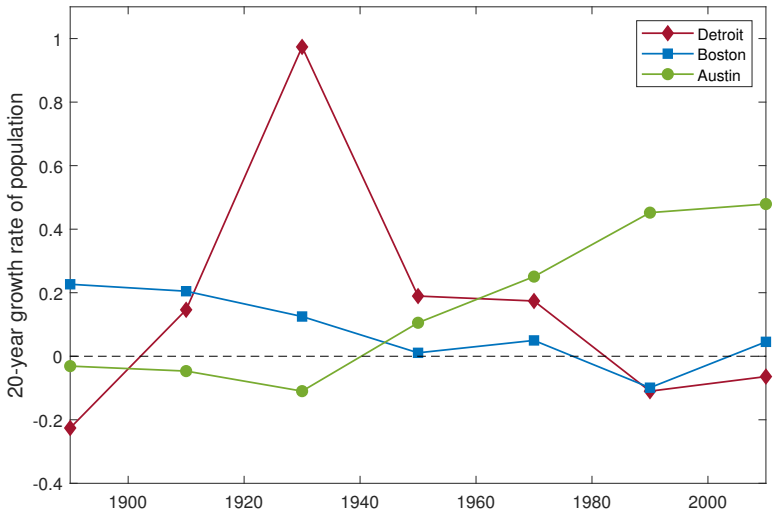
Notes: Composition of patenting output across the seven main IPC classes. Patent count for year t is constructed as the sum of patents filed between $t-10$ and $t+9$. Class names are abbreviated. The full description for each class is available at https://www.wipo.int/classifications/ipc/en/.

cities experience heterogeneous productivity gains from common technological shocks, contributing to explaining the heterogeneous historical dynamics of U.S. urban and regional growth.

The experiences of Detroit, Austin, and Boston illustrate this point. Figure 2 depicts the 20-year population growth of these three CZs since 1890, residualized with respect to Census Division-time fixed effects, which account for systematic regional variations in population growth over time. In the decades following the rise of the automobile industry around 1910, Detroit displays the highest growth rates, followed by a long-lasting decline that resulted in a steady population loss since the 1980s. Austin experienced a specular trajectory. Despite relative decline in the first half of the twentieth century, Austin has emerged as a leading innovation hub in recent decades, becoming one of the country's fastest-growing cities. Finally, Boston has maintained a significantly less volatile trajectory over the previous century, marked by periods of moderate relative growth interrupted by occasional periods of modest relative decline. As we argue in this paper, the diversification of Boston's patenting output may have made its growth trajectory less sensitive to changes in the technological landscape.[13]

[13] Glaeser (2005) discusses the causes of the slow decline of Boston between 1920 and 1980, and the subsequent re-emergence of the city, and proposes the high density of human capital as the major factor behind its resilience.

Figure 2: **City dynamics**



Notes: Residuals of a regression of 20-year growth rate of population on Census Division-time fixed effects, 1890-2010.

2.3 Technological waves and the growth and decline of cities

Figures 1 and 2 suggest that changes in the importance of technological fields may have differential effects on the growth trajectories of cities due to their pre-existing specialization across fields. While this idea is present in much of the economic geography literature, empirical support has been predominantly anecdotal, and it remains unclear to what extent these patterns are directly attributable to innovation as opposed to production activities. In this section, we show that these patterns hold systematically over the long period covered by our data and are robust to controlling for potential local confounders such as industrial composition and human capital density.

We refer to changes in the technological landscape, captured by shifts in the composition of national patenting by class-group, as "technological waves." We explore the hypothesis that cities with portfolios initially concentrated in expanding fields are better positioned to exploit new innovation opportunities, resulting in higher productivity and population growth in these cities. To capture this idea in a simple empirical setting, we construct the following measure of local exposure to technological waves:

$$Exp_{n,t} \equiv \sum_{s \in S} Share_{n,s,t-1} \times g_{-n,s,t}, \tag{1}$$

where $Share_{n,s,t-1}$ is the share of patents filed in CZ n belonging to class-group s at time $t-1$ and $g_{-n,s,t}$ is the growth rate in the share of patents of class-group s in all the other CZs between $t-1$ and t. This exposure measure is analogous to a shift-share variable (Bartik, 1991) in which the shares

correspond to the distribution of local patenting across class-groups, and the shifts to changes in the distribution of nationwide patenting across class-groups (leaving out local patenting). A city whose portfolio of patents is initially concentrated in expanding (declining) class-groups will record a positive (negative) value of $Exp_{n,t}$, reflecting a favorable (adverse) exposure to the current technological wave.

We then estimate the relationship between the growth rate of local population and our exposure measure between 1910 and 2010 by running the following regression:

$$\Delta \log (Pop_{n,t}) = \alpha \, Exp_{n,t} + \sum_{\tau=1}^{2} \beta_\tau \log (Pop_{n,t-\tau}) + \gamma \, X_{n,t} + \delta_{d(n),t} + \varepsilon_{n,t}, \tag{2}$$

where $\Delta \log (Pop_{n,t})$ is the 20-year growth rate of population in CZ n between $t-1$ and t, and $Exp_{n,t}$ denotes the exposure measure from Equation (1). The regression model includes two lags of log-population, to account for size and growth effects, such as convergence and persistence, and Census Division-time fixed effects ($\delta_{d(n),t}$), to account for the differential growth rates of CZs across space as a result of factors such as the Westward expansion or the Great Migration.[14] The term $X_{n,t}$ denotes an additional set of controls that we discuss below. We weight observations by beginning-of-the-period share of total population.[15]

Table 1 displays the results of the regression. The estimate in column 1 is positive and significant, indicating that cities with a more favorable initial exposure to the technological wave have experienced systematically greater population growth.[16] The estimated coefficient implies that an increase of one residual standard deviation in the measure of exposure is associated with an increase of 15.2% of one residual standard deviation in population growth. In column 2, we control for lagged total patenting. This control does not significantly affect the exposure measure's coefficient, suggesting that the estimated relationship is not driven by local innovation intensity.

In our analysis, we do not take a position on the factors that affect the national patenting shares by class-group. Technological waves could result from scientific and technological advancements, but also from political and environmental factors, such as regulation, trade agreements, or changes in consumer preferences. Critical to our analysis is the fact that, regardless of their source, technological waves *differentially* affect the returns to innovation in different fields and, consequently, the evolution of patenting shares across class-groups.

An alternative view is one of innovation as a byproduct of production. In this case, factors

[14] The earliest period corresponds to population growth between 1890 and 1910, and it controls for two lags of log-population (1870 and 1890). The latest period corresponds to population growth between 1990 and 2010, and it controls for two lags of log-population (1970 and 1990).

[15] In Appendix Table A.3, we show that results do not change significantly when we run unweighted regressions, or include CZ fixed effects to account for systematic differences across locations in their growth rates and their exposure measure.

[16] The relationship is visually displayed in the bin-scatter plot in Appendix Figure A.3.

that impact patenting across fields may be correlated with other industry-level shocks that drive differences in population growth across cities, confounding our interpretation. To address this concern, we construct a variable similar to the one in Equation (1), but we use employment by industry instead of patenting by class-group to compute local shares and aggregate shifts.[17] Appendix Figure A.4 plots the measures of exposure to industry shocks and to technological waves, after residualizing both variables with respect to two lags of log-population and Census Division-time fixed effects. The correlation between the two variables is positive but weak (the R^2 of the underlying regression is 0.008). This suggests substantial variation in the composition of patents across fields that is not explained by the local employment distribution across industries. The absence of a strong correlation may be attributable to several factors, including the geographical separation between innovation and production activities and the applicability of ideas within a given patent class to multiple industries.

In column 3 of Table 1, we re-estimate Equation (2) while directly controlling for industry shock exposure. This variable is a strong predictor of population growth, and its inclusion slightly reduces the estimated coefficient on the exposure to technological waves, which nevertheless remains large and significant. Finally, in column 4, we include the control for human capital density. This indicator correlates with local population growth, as documented by Glaeser and Saiz (2003). However, it has a negligible effect on the estimated coefficient of the exposure measure, suggesting that this measure does not merely reflect the availability of human capital in the city.

In summary, our empirical analysis showed a systematic correlation between local exposure to technological waves and local population growth, even after controlling for innovation intensity, industry composition, and human capital density. While this correlation can be explained through various mechanisms, such as irreversible investments in specialized physical or human capital, the remainder of the paper examines a specific hypothesis that rationalizes this relationship: the existence of frictions to the diffusion of ideas across space and fields of knowledge, which prevent cities from reallocating their resources optimally to capitalize on technological waves. Cities with an innovation portfolio skewed toward expanding fields are better positioned to take advantage of new innovation opportunities and will become more attractive to workers and firms.

In the following sub-section, we use patent citation data to provide suggestive evidence of the existence of geographical and technological frictions to idea diffusion. We then develop a model to structurally quantify the contribution of these frictions to the empirical relationship between exposure to technological waves and population growth documented in this section.

[17] See Appendix B for details on the construction of the data on employment by industry at the commuting zone level. Industries correspond to the 12 main industries in the 1950 Census Bureau industrial classification system.

Table 1: **Technological waves and city growth**

	Growth rate of population			
	(1)	(2)	(3)	(4)
Exposure to tech. waves, $Exp_{n,t}$	0.429***	0.383***	0.350***	0.341***
	(0.083)	(0.087)	(0.084)	(0.099)
Log-total patents		0.033***	0.008	0.006
		(0.011)	(0.013)	(0.014)
Industry composition			0.570***	0.573***
			(0.113)	(0.113)
Human capital (ranking)				0.020
				(0.066)
Log-population (lags 1 and 2)	Yes	Yes	Yes	Yes
Fixed effects	CD×T	CD×T	CD×T	CD×T
# Obs.	2,910	2,881	2,855	2,852
R^2	0.512	0.519	0.545	0.544

Notes: CZ-time-level regression, 1910-2010, weighted by share of population at the beginning of the period. Dependent variable defined as growth rate of population over 20 years. "CD×T" denotes Census Division-time fixed effects. Standard errors clustered at the CZ and the Census Division-time level in parenthesis. ***$p < 0.01$.

2.4 Evidence on frictions to knowledge diffusion

The fact that knowledge diffusion is highly localized has been well documented in the literature on the geography of innovation. Within this literature, a rich body of work, starting with Jaffe et al. (1993), has provided evidence of this localization by studying the spatial patterns of patent citations (Murata et al., 2014; Kerr and Kominers, 2015). Our citation data confirm this evidence of localization, which does not appear to diminish over time. Citations to the same CZ account for 15.6% of all citations for patents filed since 1940.[18] This proportion of citations is 6.24 times the background probability (2.5%) of observing a citation to the same CZ if citations were distributed randomly and proportionally to each city's overall share of patents. When we divide the sample into an early (1940–1979) and late (1980 and later) subsamples, we find that, despite the dramatic decline in communication costs, this evidence of localization has, if anything, strengthened over time. In the early subsample, the proportion of citations to the same CZ is 4.34 times greater than

[18]Patent citations are not consistently available in the earlier decades. Thus, when considering citations, we restrict the sample to all the patents filed since 1940. A separate section containing referenced patents was formally introduced in patent documents only in 1947. In constructing these statistics, we only consider citations to and from commuting zones included in our sample, and weight each citation by the inverse of the total number of citations given by the citing patent. By doing so, each citing patent has a weight of one in our sample.

the background probability and in the late subsample, it is 6.60 times greater.[19]

Analogously, we find substantial evidence of localization of patent citations in the technological space. Appendix Figure A.5 displays a heatmap of the empirical probability that a citation from each class-group on the vertical axis is directed toward each class-group on the horizontal axis. The concentration of probabilities along the diagonal of the heatmap indicates a high degree of technological localization in the diffusion of ideas. Appendix Figure A.6 provides separate heatmaps for the early subsample (1940–1979) and the late subsample (1980 and later), showing that evidence of technological localization remains strong over time.

3 Stylized model and theoretical predictions

Motivated by our empirical findings, in this section we develop a model that incorporates endogenous growth via frictional idea diffusion into a spatial equilibrium framework. We start by presenting a stylized version of the model, which allows us to solve for the equilibrium in closed form and to derive sharp theoretical predictions to gain insight into the model's mechanics. In Section 4, we extend the model to incorporate relevant features from a quantitative standpoint, and use this extended version for the quantitative analysis of Section 5.

In the model, agents decide where to migrate and in which sector to specialize after forming expectations about their lifetime productivity. This productivity is determined through a decision of whether to imitate an idea from the local distribution, or to innovate upon an idea drawn from the distribution of any location and sector in the economy. The applicability of an idea for innovation is constrained by diffusion frictions—so that knowledge acquired from geographically and technologically closer sources can be converted into new inventions more efficiently—and is affected by exogenous technological wave shocks that change the effectiveness with which ideas from specific fields can be used as inputs in the innovation process. The presence of frictions to diffusion implies that common technological wave shocks induce heterogeneous responses in local growth, providing a rationale for the reduced-form relationship documented in Section 2.

3.1 Setting

The economy comprises a finite set N of locations and a finite set S of sectors. Time is discrete and indexed by t. At each point in time, a mass L_t of individuals populate the economy. In what follows, N and S denote the sets of locations and sectors and their cardinality, and L_t denotes the set of individuals at time t and its mass.

[19] The observed probabilities are 13.9% in the early and 16.5% in the late sub-samples, whereas the background probabilities are 3.2% and 2.5%, respectively.

3.1.1 Preferences, endowments, and demographics

In each period, a new generation is born in the location of their parents. Individuals live for one period and have f_t children who enter the economy in the following period.

At the beginning of the period, newborn agents choose where to migrate and in which sector to specialize to maximize expected utility, subject to idiosyncratic utility draws that affect the desirability of each location-sector. Each individual $i \in L_t$ receives a complete set of stochastic utility draws, one for each location-sector in the economy:

$$\mathbf{x}_i = \{x_{n,s,i}\}_{(n,s)\in N\times S}.$$

Each value $x_{n,s,i}$ is a random draw from a Fréchet distribution with shape parameter $\zeta > 1$. Individual i migrating to location n and specializing in sector s obtains utility

$$U(u_n, x_{n,s,i}, c_{n,s,i,t}) = u_n\, x_{n,s,i}\, c_{n,s,i,t}, \tag{3}$$

where u_n is the level of amenities in city n and $c_{n,s,i,t}$ denotes consumption of the final good.[20] Because individual labor productivity will be stochastic, agents will face uncertainty regarding their consumption $c_{n,s,i,t}$ when selecting a location-sector (i.e., after the idiosyncratic utility draws $\mathbf{x}_i$ are realized). Consequently, they will select the location-sector that offers the greatest expected utility. We return to this point in Section 3.1.4.

3.1.2 Production technology, consumption, and total output

Each agent i inelastically supplies one unit of labor with productivity $q_{n,s,i,t}$ to produce a tradable and homogeneous final good, whose price in each period is normalized to one. Consequently, the wage of agent i is simply their productivity $q_{n,s,i,t}$. Given that agents live for one period, their consumption of final good is given by their own production:

$$c_{n,s,i,t} = q_{n,s,i,t}.$$

Total output in the economy is given by a linear aggregator of individual productivity across all locations and sectors:

$$Y_t = \sum_{n\in N}\sum_{s\in S} L_{n,s,t}\mathbb{E}[q_{n,s,\cdot,t}], \tag{4}$$

where $L_{n,s,t}$ denotes the mass of agents in location-sector (n,s) and $\mathbb{E}[q_{n,s,\cdot,t}]$ denotes their average

[20] Note that there are no moving costs across locations and amenities are assumed to be exogenous and time-invariant. These assumptions will be relaxed in Section 4 by allowing for moving costs and for amenities with an exogenous time-varying component and an endogenous component capturing local congestion forces.

productivity.[21]

3.1.3 Imitation, innovation, and knowledge diffusion

Individual productivity is endogenously determined through a choice between imitating an available technology or enhancing an existing idea by generating an innovation. The quality of ideas in the choice set of each agent is stochastic, and its distribution varies by location-sectors. After agent i has chosen their location-sector (n, s), they receive a complete set of idiosyncratic, independently distributed draws:

$$\mathbf{z}_{n,s,i} = \left\{ z^l_{n,s,i}, \left\{ z^x_{m,r,i} \right\}_{m,r \in N \times S} \right\}. \tag{5}$$

The first term, $z^l_{n,s,i}$, is a random draw from the distribution of productivity in location-sector (n, s) at time $t-1$, whose cumulative distribution is denoted by $F_{n,s,t-1}(q)$. This draw can be interpreted as knowledge acquired from teachers, mentors, or managers, which can be imitated and directly applied to production. The second set of terms, $\{z^x_{m,r,i}\}_{m,r \in N \times S}$, is a full vector of random draws from each productivity distribution in all locations and sectors (including local ones, $m = n$ and $r = s$) at time $t-1$, with corresponding cumulative distributions $\{F_{m,r,t-1}(q)\}_{m,r \in N \times S}$. These draws can be interpreted as knowledge acquired through various transmission channels, such as books, radio, television, and the internet, or through informal interactions with local or non-local individuals. Although these ideas cannot be directly adopted in production, they can serve as an input for innovation, as we describe below.

After observing the full set of draws, $\mathbf{z}_{n,s,i}$, the agent must select one of these draws, which will determine their lifetime productivity. If the agent adopts the local draw, $z^l_{n,s,i}$, their lifetime productivity will be

$$q_{n,s,i,t} = z^l_{n,s,i}.$$

Since the production technology used by the agent was already available in the previous generation, we refer to this choice as "imitation."

Alternatively, if the agent chooses one of the external draws, $z^x_{m,r,i}$, their lifetime productivity will be

$$q_{n,s,i,t} = \frac{\epsilon_{n,s,t}\, \alpha_{r,t}\, z^x_{m,r,i}}{d_{(m,r)\to(n,s)}}. \tag{6}$$

Since in this case the agent has developed a new production technology that was not previously

[21] To focus on the interplay between idea diffusion and local growth, the model abstracts from other economic forces such as inter-regional trade and capital accumulation. While these forces could interact with the model's mechanism by shaping local income by sector, in Section 4.3 we show that the model can closely account for the empirical variation in income across location-sectors even abstracting from those forces, suggesting that including them explicitly is unlikely to significantly alter the results. Since inter-regional trade can directly impact knowledge diffusion, in estimating frictions we work with a flexible parameterization that includes a direct effect of geographical distance, encapsulating the effects of trade and other diffusion channels related to distance on idea transmission.

available, we refer to this choice as "innovation." In Equation (6), the term $d_{(m,r)\to(n,s)} \geq 1$ captures knowledge transmission frictions between the location and sector of origin of the idea, (m, r), and of destination, (n, s). For now, we treat this term as exogenous and time-invariant, and we do not impose a specific structure on it. In the quantitative analysis, this term will be parametrized flexibly to capture a wide range of channels affecting the efficacy of knowledge diffusion.[22] The term $\alpha_{r,t}$ captures technological waves shocks. It represents the importance of sector r in the innovation landscape at time t: the greater its value, the more effectively knowledge in sector r can be translated into innovations for all sectors.[23] Finally, $\epsilon_{n,s,t}$ is a structural residual that captures the current efficacy of innovation in (n, s) and is shared by all innovators in that location-sector. It accounts for all residual factors that affect the local sector's efficiency of innovation, such as the opening of research facilities.

In sum, agent i in location-sector (n, s) chooses whether to imitate or innovate to maximize their lifetime productivity:

$$q_{n,s,i,t} = \max \left\{ z^l_{n,s,i},\ \max \left\{ \frac{\epsilon_{n,s,t}\, \alpha_{r,t}\, z^x_{m,r,i}}{d_{(m,r)\to(n,s)}} \right\}_{m,r \in N \times S} \right\} \tag{7}$$

Equation (7) shows how this process can be divided into two steps. First, agent i chooses the best available innovative idea. Then, agent i compares this idea with the imitation draw, $z^l_{n,s,i}$, and picks the one that yields higher productivity.

The following assumption, which we maintain throughout the paper, plays a vital role in keeping the theory tractable:

Assumption A1. *The initial productivity distribution $F_{n,s,0}(q)$ in all location-sectors (n, s) is Fréchet with shape parameter $\theta > 1$ and scale parameter $\lambda_{n,s,0} > 0$:*

$$F_{n,s,0}(q) = e^{-\lambda_{n,s,0} q^{-\theta}}. \tag{8}$$

The maximum over a set of Fréchet distributed random variables with common shape parameter is itself Fréchet with the same shape parameter. Hence, under Assumption A1, individual optimal choice between imitation and innovation (Equation 7) implies that productivity at any time $t \geq 0$ is distributed Fréchet with shape parameter θ and with scale parameter evolving according to the

[22] The parametrization will account for channels such as proximity in the migration network (Cai et al., 2022b), barriers to applicability and intelligibility of ideas across technological areas (Hayes, 1992), and, in reduced form, other channels of diffusion that are influenced by geographical distance, such as trade linkages (Buera and Oberfield, 2020) and frequency of business trips between locations (Catalini et al., 2020; Pauly and Stipanicic, 2022).

[23] Note that $\alpha_{r,t}$ is inherent to the sector of origin r at time t, and is unrelated to agent i's location-sector (n, s).

following law of motion:

$$\lambda_{n,s,t} = \underbrace{\lambda_{n,s,t-1}}_{\text{Imitation}} + \underbrace{\sum_{m \in N} \sum_{r \in S} \lambda_{m,r,t-1} \left(\frac{\epsilon_{n,s,t}\, \alpha_{r,t}}{d_{(m,r)\to(n,s)}} \right)^{\theta}}_{\text{Innovation}}. \tag{9}$$

Equation (9) is crucial to our theory, because it describes the dynamics of productivity across location-sectors.[24] The scale parameter of the knowledge distribution in the new generation, $\lambda_{n,s,t}$, is equal to the scale parameter of the previous generation augmented by a second term that captures inventive activities. This second term is the sum of scale parameters across all location-sectors, weighted by their applicability to location-sector (n, s). The applicability term includes the importance of each field of knowledge, $\alpha_{r,t}$, the local effectiveness of innovation, $\epsilon_{n,s,t}$, and is discounted by technological and geographical frictions between location-sector pairs, $d_{(m,r)\to(n,s)}$. Equation (9) illustrates that innovation is the only source of growth in the knowledge frontier, while imitation is what carries forward knowledge developed in previous periods.

The scale parameter of the productivity distribution summarizes the knowledge stock in each location-sector. Specifically, because the productivity distribution in (n, s) at time t is Fréchet with shape θ and scale $\lambda_{n,s,t}$, the local average productivity is

$$\mathbb{E}\left[q_{n,s,\cdot,t}\right] = \Gamma\left(1 - \frac{1}{\theta}\right) \lambda_{n,s,t}^{\frac{1}{\theta}}, \tag{10}$$

where $\Gamma(\cdot)$ denotes the gamma function.

The process of knowledge diffusion described by Equation (7) and the Fréchet assumption A1 imply that the probability that an innovator in location-sector (n, s) builds upon an idea from any location-sector (m, r) at time t can be expressed as follows:

$$\eta_{(m,r)\to(n,s),t} = \frac{\lambda_{m,r,t-1} \left(\frac{\alpha_{r,t}}{d_{(m,r)\to(n,s)}} \right)^{\theta}}{\sum_{l \in N} \sum_{p \in S} \lambda_{l,p,t-1} \left(\frac{\alpha_{p,t}}{d_{(l,p)\to(n,s)}} \right)^{\theta}}. \tag{11}$$

This expression shows that the likelihood that innovators rely on ideas from a given origin is increasing in the average quality of ideas in the origin location-sector ($\lambda_{m,r,t-1}$) and the importance of the sector of origin ($\alpha_{r,t}$), and decreasing in the geographical and technological frictions between the location-sectors of origin and destination ($d_{(m,r)\to(n,s)}$).

[24] This result follows from the max-stability property, which states that the maximum of a vector of independent random draws $\{a_k\, z_k\}_{k=1}^{K}$, with $a_k > 0$ and z_k distributed Fréchet with scale λ_k and common shape θ is itself distributed Fréchet with scale $\sum_{k=1}^{K} \lambda_k a_k^{\theta}$ and shape θ. Tractability can be maintained without assuming independence, as in Lind and Ramondo (2019).

3.1.4 Migration and occupational choice

At the beginning of period t, agents in the new generation observe their idiosyncratic utility draws ($\mathbf{x}_i$) as well as sectoral and local shocks ($\alpha_{r,t}$ and $\epsilon_{n,s,t}$ for all $n \in N$ and $r, s \in S$), but do not know the idiosyncratic idea draws ($\mathbf{z}_{n,s,i}$) they will receive in their location-sector of choice. They have to form expectations about their future wages (determined by the idea draws) before making their migration and occupational decisions. Agent i moving to location-sector (n, s) has expected utility equal to

$$\mathbb{E}\left[U(u_n, x_{n,s,i}, q_{n,s,i,t})\right] = u_n \, x_{n,s,i} \, \mathbb{E}\left[q_{n,s,\cdot,t}\right]. \tag{12}$$

Combining Equations (10) and (12), the probability that any newborn individual selects location-sector (n, s) is

$$\pi_{n,s,t} = \frac{\left(u_n \, \lambda_{n,s,t}^{\frac{1}{\theta}}\right)^{\zeta}}{\sum_{m \in N} \sum_{r \in S} \left(u_m \, \lambda_{m,r,t}^{\frac{1}{\theta}}\right)^{\zeta}}. \tag{13}$$

This expression is intuitive: the probability of choosing location-sector (n, s) is increasing in its expected productivity (which is proportional to $\lambda_{n,s,t}^{\frac{1}{\theta}}$) and its appeal due to amenities (u_n) relative to the average across location-sectors appearing in the denominator. The mass of agents in location-sector (n, s) at time t is equal to the probability of choosing (n, s) times the population size:

$$L_{n,s,t} \equiv \pi_{n,s,t} \, L_{t-1} \, f_t. \tag{14}$$

3.1.5 Equilibrium

An equilibrium in this economy is given by a path of population and productivity for each location and sector that, given the initial conditions and a path of the exogenous variables, is consistent with individual optimal decisions on lifetime productivity (Equation 7) as well as migration and occupation (Equation 13). We formally define the equilibrium in Appendix C. It is straightforward to show that all equilibrium conditions have a unique and explicit solution. Therefore, a unique equilibrium exists and can be expressed in closed form for any initial conditions and exogenous variable path.

3.1.6 Existence and uniqueness of a BGP

We define a BGP as an equilibrium in which sectoral importance $\alpha_{r,t}$ and structural residuals $\epsilon_{n,s,t}$ are constant over time, and scale parameters $\lambda_{n,s,t}$ grow at the same rate for all location-sectors (n, s). Using Equation (13), these conditions also imply that migration and occupational choices (and, as a result, the distribution of individuals across locations and sectors) are stable over time.

Notice that Equation (9) can be rewritten as a difference equation in matrix form:

$$\vec{\lambda}_{t+1} = A_t \vec{\lambda}_t, \tag{15}$$

where $\vec{\lambda}_t$ is a $N \times S$ vector of all scale parameters $\lambda_{n,s,t}$, and A_t is the $(N \times S)^2$ diffusion matrix implied by Equation (9). In BGP, the matrix A_t is constant, and we denote it by A^* (in what follows, we use stars to denote variables at their BGP value).

It follows from Equation (15) that in BGP $\vec{\lambda}_t$ must be an eigenvector of A^*, with the corresponding eigenvalue equal to its gross growth rate $1 + g^*_\lambda$. The Perron-Frobenius theorem states that A^* has a unique positive eigenvector (and corresponding eigenvalue), provided that every entry in A^* is positive. A sufficient condition for A^* to contain only positive entries is that the frictions to knowledge diffusion $d_{(m,r)\to(n,s)}$ are strictly positive and finite for each combination of location-sector pairs, which we have assumed. This proves the following:

Proposition 1. *Let $1 \leq d_{(m,r)\to(n,s)} < +\infty$ for all $(m,r),(n,s) \in (N \times S)^2$. Then, for each set of constant sectoral importance $\{\alpha^*_r\}_{r\in S}$ and structural residuals $\{\epsilon^*_{n,s}\}_{(n,s)\in N\times S}$, there exists a unique BGP in which $\{\lambda_{n,s,t}\}_{(n,s)\in N\times S, t\geq 0}$ grow at constant rate, g^*_λ, with $g^*_\lambda > 0$. The gross growth rate $(1 + g^*_\lambda)$ is given by the unique largest eigenvalue of A^* (the Perron-Frobenius eigenvalue), and the normalized scale parameters $\{\lambda_{n,s,t}/(1 + g^*_\lambda)^t\}_{(n,s)\in N\times S, t\geq 0}$ correspond to the associated right eigenvector of A^*. Along the BGP, the gross growth rate of output per worker is given by $(1+g^*_\lambda)^{\frac{1}{\theta}}$.*

As stated in the proposition, along the BGP, scale parameters grow at the same rate across location-sectors. However, location-sectors have different scale parameters because the entries of the eigenvector associated with the Perron-Frobenius eigenvalue are generally distinct. In BGP the scale parameter relative to the mean, $\tilde{\lambda}_{n,s} = \frac{\lambda_{n,s}}{\mathbb{E}[\lambda_{\cdot,\cdot}]}$, is determined by the following equation, which holds true for each location-sector (n,s):

$$\tilde{\lambda}^*_{n,s} = \frac{(\epsilon^*_{n,s})^\theta}{g^*_\lambda} \sum_{m\in N} \sum_{r\in S} \tilde{\lambda}^*_{m,r} \left(\frac{\alpha^*_r}{d_{(m,r)\to(n,s)}} \right)^\theta. \tag{16}$$

This equation shows that the stationary value of the scale parameters, and hence average productivity, depends on the matrix of diffusion frictions across location-sectors, $d_{(m,r)\to(n,s)}$, as well as local and sectoral characteristics, α^*_r and $\epsilon^*_{n,s}$, given a growth rate g^*_λ.[25] The gross growth rate of output per worker, $(1 + g^*_\lambda)^{\frac{1}{\theta}}$, follows from Equation (10).

[25] Another work that studies the BGP properties of an endogenous growth model with spillovers across multiple sectors is Huang and Zenou (2020). While the setting for idea diffusion in Huang and Zenou (2020) differs from ours, the Perron-Frobenius theorem is central to establishing the existence and uniqueness of a BGP equilibrium in both models.

3.2 Log-linearized model dynamics

Although the BGP is a useful benchmark, we are ultimately interested in the heterogeneous response of cities to technological wave shocks. By design, the BGP analysis holds technological waves constant ($\alpha_{r,t} \equiv \alpha_r^*$) and, as a result, the relative size of cities along the BGP does not change. We now study the model's dynamics by log-linearizing the equilibrium conditions around the BGP. This allows us to derive intuitive characterizations of what drives city growth in response to technological waves. For this purpose, we assume that the economy is in a BGP at time $t-1$. At time t, the economy is hit by technological wave shocks $\{\hat{\alpha}_{r,t}\}_{r\in S}$, where hats represent log-deviations from BGP values.

First, we consider the dynamics of the scale parameter of the local distribution of productivity, $\lambda_{n,s,t}$. Log-linearizing Equation (9) around the BGP yields

$$\hat{\lambda}_{n,s,t} = \frac{\theta(\epsilon_{n,s}^*)^\theta}{1+g_\lambda^*} \sum_{m,r} \left(\frac{\lambda_{m,r}}{\lambda_{n,s}}\right)^* \left(\frac{\alpha_r^*}{d_{(m,r)\to(n,s)}}\right)^\theta \hat{\alpha}_{r,t}. \tag{17}$$

By multiplying and dividing the right-hand side of (17) by g_λ^*, and using (11) and (16), we derive the following proposition that links changes in local productivity to technological wave shocks.

Proposition 2. *The log deviation of the scale parameter of the productivity distribution of (n,s) from the BGP, $\hat{\lambda}_{n,s,t}$, is equal to the sum of the shock to r, $\hat{\alpha}_{r,t}$, over all sectors $r \in S$, weighted by the reliance of innovation in (n,s) on ideas from sector r, $\eta^*_{r\to(n,s)} \equiv \sum_{m\in N} \eta^*_{(m,r)\to(n,s)}$:*

$$\hat{\lambda}_{n,s,t} = \frac{\theta g_\lambda^*}{1+g_\lambda^*} \sum_{r\in S} \eta^*_{r\to(n,s)} \hat{\alpha}_{r,t}. \tag{18}$$

Proposition 2 implies that the sensitivity of local productivity to shocks to any given sector, $\hat{\alpha}_{r,t}$, is increasing in the probability of drawing ideas from that sector to innovate, $\eta^*_{r\to(n,s)}$. Due to diffusion frictions, this reliance depends on the geographical and technological proximity of r to (n,s). Quantitatively, given the localized nature of diffusion, the local stock of knowledge in the same sector, $\lambda_{n,r}$, will play a decisive role in determining this reliance term. Since this stock of knowledge is tightly linked to the local employment share in the same sector, this share will also be a key determinant of the sensitivity of local productivity to shocks to r.

Using the previous result on the evolution of the scale parameters, we can next characterize the population dynamics in location n. Combining Equation (13) with the definition $\pi_{n,t} \equiv \sum_{s\in S} \pi_{n,s,t}$ and log-linearizing the resulting expression for small deviations of $\{\lambda_{m,s,t}\}_{m,s\in N\times S}$ from their BGP values yields

$$\hat{\pi}_{n,t} = \frac{\zeta}{\theta} \sum_{s\in S} \left\{ (1-\pi_n^*)\pi^*_{s|n} \hat{\lambda}_{n,s,t} - \sum_{m\neq n} \pi^*_{m,s} \hat{\lambda}_{m,s,t} \right\}, \tag{19}$$

where $\pi^*_{s|n}$ denotes the probability of being employed in sector s conditional on living in location n.[26] Equation (19) implies that a location grows relative to the rest of the economy if and only if changes in local sectoral productivities, weighted by the incidence of each sector in the city, are larger than the average corresponding changes for the rest of the economy:

$$\hat{\pi}_{n,t} > 0 \iff \sum_{s \in S} \pi^*_{s|n} \hat{\lambda}_{n,s,t} > \sum_{s \in S} \sum_{m \neq n} \frac{\pi^*_{m,s}}{1 - \pi^*_n} \hat{\lambda}_{m,s,t}.$$

Combining Equations (18) and (19), we characterize population dynamics as a function of BGP values and technological wave shocks:

Proposition 3. *The log-change in the population shares of location n, $\hat{\pi}_{n,t}$, depends on technological wave shocks as follows:*

$$\hat{\pi}_{n,t} = \frac{\zeta g^*_\lambda}{1 + g^*_\lambda} \sum_{r \in S} \sum_{s \in S} \left\{ (1 - \pi^*_n) \pi^*_{s|n} \eta^*_{r \to (n,s)} - \sum_{m \neq n} \pi^*_{m,s} \eta^*_{r \to (m,s)} \right\} \hat{\alpha}_{r,t}. \tag{20}$$

To interpret Equation (20) and better illustrate the economic mechanism at play, we first consider the limit case in which knowledge flows across sectors are of second-order importance relative to flows within sectors, and individual cities are of negligible size relative to the rest of the economy. In particular, we impose the following assumption:

Assumption A2. *(for illustration purposes only)*

1. *Frictions to knowledge diffusion across sectors are large enough that, effectively, knowledge flows occur only within sectors, that is, $\eta^*_{s \to (n,s)} \approx 1$ for all $s \in S$.*

2. *The size of any given city is negligible relative to the entire economy, that is, $\sum_{m \neq n} \pi^*_m \approx 1$, for all $n \in N$.*

Rewriting Equation (20) under Assumption A2, we obtain

$$\hat{\pi}_{n,t} \propto b_t + \sum_{s \in S} \pi^*_{s|n} \hat{\alpha}_{s,t}, \tag{21}$$

The term $b_t \equiv -\sum_{s \in S} \pi^*_{\cdot,s} \hat{\alpha}_{s,t}$, with $\pi^*_{\cdot,s}$ denoting the share of the national population employed in sector s, is a common term across all locations. Equation (21) shows that, under Assumption A2,

[26] Equation (19) holds in the absence of migration frictions across locations. We derive the corresponding expression when bilateral migration costs are introduced in our setting in Appendix D. Intuitively, in the presence of migration frictions, local population growth is determined by local productivity growth relative to productivity growth in other locations, where locations with lower migration costs have more weight than locations with higher migration costs. Equation (D.3), combined with Equation (D.4), is equivalent to the expression for changes in regional labor supply obtained by Borusyak et al. (2022). A similar mechanism also emerges in Schubert (2021).

cities' differential patterns of population growth depend on a weighted average of sectoral shocks, with the weights corresponding to the city's pre-existing sectoral shares.[27] Equation (21) thus provides a rationale for the shift-share functional form used to measure exposure to technological waves in the reduced-form analysis of Section 2.3.[28]

Consider now population dynamics when we allow for knowledge flows across fields, while maintaining the assumption of every city being of negligible size (in other words, we drop Assumption A2.1 and only impose A2.2). In this case, Equation (20) can be rewritten as

$$\hat{\pi}_{n,t} \propto b_t + \sum_{s \in S} \sum_{r \in S} \pi^*_{s|n} \eta^*_{r \to (n,s)} \hat{\alpha}_{r,t}, \tag{23}$$

where b_t is again a common term across all locations. In this case, because of geographical frictions to idea diffusion, cities display different degrees of reliance of local innovation on ideas from each sector (as captured by $\eta^*_{r \to (n,s)}$). This implies that in cities where expanding (shrinking) fields are more prevalent, productivity growth in *all* sectors will be larger (smaller), thereby amplifying the "shift-share" effect in Equation (21). In other words, localized knowledge flows *across* fields amplify fluctuations in productivity growth and, as illustrated by Equation (23), fluctuations in population dynamics in response to technological wave shocks.

Finally, relaxing also the second part of Assumption A2 that all cities are of negligible size, yields the full population dynamics in Equation (20). This equation highlights that the effect of technological waves on local growth depends on sectoral shares ($\pi^*_{s|n}$) and reliance terms ($\eta^*_{r \to (n,s)}$), as in Equation (23), but only *relative* to the rest of the economy. When individual cities are of non-negligible size, the rest of the economy is location-specific, and therefore the entire geography of the economy matters. Equation (20) takes this fact into account.

Taking stock Proposition 3 shows that frictions to idea diffusion imply rich and heterogeneous effects of technological waves on local growth. Frictions across sectors imply that technological wave shocks affect local population dynamics through local employment shares, as illustrated by Equation (21). Frictions across geographical areas amplify this effect by generating variation in the local reliance of innovators on different fields, as shown in Equation (23). In the remainder of the paper, we assess the quantitative importance of these channels.

[27] In this sense, Equation (21) can be interpreted as a micro-foundation for commonly used Bartik shocks, that we derive within a fully specified theory of migration and local productivity growth.

[28] Equation (21) also implies that under Assumption A2, n grows (shrinks) if and only if the average local exposure to the technological wave is larger (smaller) than the average exposure for the economy:

$$\hat{\pi}_{n,t} > 0 \iff \sum_{s \in S} \pi^*_{s|n} \hat{\alpha}_{s,t} > \sum_{s \in S} \pi^*_{\cdot,s} \hat{\alpha}_{s,t}. \tag{22}$$

4 Quantitative model and calibration

This section describes how we extend the model by introducing potentially relevant features from a quantitative standpoint and how we bring the model to the data.

4.1 Extended model for quantitative analysis

We extend the model along three dimensions. First, we allow individuals to live for multiple periods. Second, we introduce moving costs that are increasing with geographical distance. Third, we allow for endogenous congestion forces by relaxing the assumption that amenities are exogenous and time-invariant. Existing research contends that these channels are quantitatively significant (see, among others, Diamond, 2016 and Allen and Donaldson, 2022). We show here that these channels can be incorporated into our framework without incurring a prohibitive loss of tractability. Appendix D derives the log-linearized dynamics within this richer framework, and shows that the same intuitions obtained in the simple model of Section 3 are also present in this expanded framework.

In the extended model, individuals live for three periods ("childhood", "young adults", and "old adults"). Every child is born where their parents reside. After childhood, agents select the location n to which they will migrate and the sector s in which they will specialize. This decision is irreversible, so every agent will spend their youth and old age in the same location-sector. Following migration, every young adult will have f_t children. The productivity of agent i who is young at time t and old at time $t+1$ is denoted by $q^y_{n,s,i,t}$ and $q^o_{n,s,i,t+1}$, respectively. Agents are endowed with one unit of inelastically-supplied labor per period.

Migration and occupational decisions maximize expected lifetime utility, subject to migration costs and idiosyncratic utility draws. The lifetime utility of individual i born in location m and selecting location-sector (n,s) is given by:

$$U_{m\to n}(u_{n,t}, x_{n,s,i}, c^y_{n,s,i,t}, c^o_{n,s,i,t+1}) = u_{n,t}\,\frac{x_{n,s,i}\,(c^y_{n,s,i,t})^{\beta}\,(c^o_{n,s,i,t+1})^{1-\beta}}{\mu_{m\to n}}, \tag{24}$$

where $u_{n,t}$ is the level of amenities in city n at time t, $x_{n,s,i}$ is an idiosyncratic utility draw from a Fréchet distribution with shape parameter $\zeta > 1$, $\mu_{m\to n}$ represents migration costs of moving from m to n, $c^y_{n,s,i,t}$ and $c^o_{n,s,i,t+1}$ denote consumption in the youth and the old period, and $\beta \in (0,1)$ is the weight of youth consumption in lifetime utility. There are no markets to smooth consumption over time and generations. Therefore, in every period, individual consumption is equal to production.

Total output in the economy is given by a linear aggregator of individual productivities across

all locations and sectors:

$$Y_t = \sum_{n \in N} \sum_{s \in S} \left(L^y_{n,s,t} \mathbb{E}[q^y_{n,s,\cdot,t}] + L^o_{n,s,t} \mathbb{E}[q^o_{n,s,\cdot,t}] \right),$$

where $\mathbb{E}[q^y_{n,s,\cdot,t}]$ ($\mathbb{E}[q^o_{n,s,\cdot,t}]$) denotes the average productivity of young (old) agents in location-sector (n, s), and $L^y_{n,s,t}$ ($L^o_{n,s,t}$) denotes the corresponding mass of individuals.

Young adults undergo the imitation or innovation process outlined in Section 3.1.3 to determine their youth productivity $q^y_{n,s,i,t}$ upon moving. Note that, as in the stylized model, the relevant variables ($\epsilon_{n,s,t}$ and $\alpha_{n,s,t}$) are known at the time of the migration and occupational choice. Under Assumption A1, the local distribution of productivity among young agents remains Fréchet and the corresponding scale parameter, $\lambda_{n,s,t}$, follows the law of motion in Equation (9). Old adults retain the same youth-period idiosyncratic productivity rescaled by a factor A^o which captures the experience productivity premium enjoyed by old relative to young agents:

$$q^o_{n,s,i,t+1} = A^o q^y_{n,s,i,t}.$$

As commonly assumed in the quantitative economic geography literature, local amenities are an isoelastic function of local population:

$$u_{n,t} = \bar{u}_{n,t} \, L^{\omega}_{n,t},$$

where $\bar{u}_{n,t}$ is the exogenous time-varying component of local amenities, and ω is the elasticity of local amenities to population, that can account for both congestion ($\omega < 0$) and agglomeration ($\omega > 0$) forces.

The equilibrium is defined in an analogous way as in the stylized model in Section 3, and the details are relegated to Appendix D.

4.2 Calibration

The model has a recursive structure that enables us to sequentially calibrate parameters and unobserved variables by making transparent assumptions on the mapping of model's objects to data on population, income, and patenting. As in the empirical analysis, we set the model period to 20 years, let N represent the set of 1990 CZs that accounted for at least 0.01% of the total population in each decade since 1890, and define sectors as the 11 class-groups detailed in Appendix Table A.1. We assume the economy is on a BGP in 1890 and, given the initial conditions and the calibrated path of exogenous local and aggregate shocks, it evolves endogenously until 2010, the last period of our sample.

Our calibration proceeds in four steps. In the first step, we infer moving costs $\mu_{m\to n}$ by deriving and estimating a gravity equation for migration flows. In the second step, we recover the path of local productivities $\lambda_{n,s,t}$, time-varying local amenities $u_{n,t}$, and aggregate fertility f_t, while simultaneously determining the time-invariant parameters θ (the Fréchet shape of the productivity distribution) and A^o (the experience premium). In the third step, we estimate frictions to knowledge diffusion $d_{(m,r)\to(n,s),t}$ by deriving and estimating a gravity equation for idea flows using patent citation data. In the fourth and final step, we recover technological wave shocks $\alpha_{s,t}$ and structural residuals $\epsilon_{n,s,t}$.

4.2.1 First Step: Gravity equation for migration flows

We assume migration costs between each pair of locations, $\mu_{m\to n}$, to be an exponential function of a constant term (capturing a fixed cost of moving) and the geographical distance in 1,000 km between m and n, $D_{m,n}$:

$$\mu_{m\to n} = e^{\bar{\mu}^0 \mathbf{1}_{m\neq n} + \bar{\mu}^D D_{m,n}}. \tag{25}$$

In Appendix E.1 we show that the probability that an individual born in m migrates to n, $\pi_{m\to n,t}$, can be written in gravity form as

$$\log(\pi_{m\to n,t}) = \psi^0_{m,t} + \psi^1_{n,t} - \zeta\bar{\mu}^0 \mathbf{1}_{m\neq n} - \zeta\bar{\mu}^D D_{m,n}, \tag{26}$$

where $\psi^0_{m,t}$ and $\psi^1_{n,t}$ denote origin-time and destination-time fixed effects, respectively.

We estimate Equation (26) via Poisson Pseudo Maximum Likelihood (PPML).[29] For the estimation of (26) and for the other time-invariant parameters pertaining to migration, population, and income, we utilize IPUMS data from 1990, which is the most recent period in the model for which comprehensive demographic data are available. We obtain a statistically significant estimate of the composite semi-elasticity of migration to distance, $\zeta\bar{\mu}^D$, of 2.81 (see Appendix Table E.5). This estimate indicates that a 100 km increase in distance reduces the migration probability by about 24%.

In our baseline, we set ζ, the elasticity of migration (or sectoral mobility) with respect to average income, equal to 4, which is the benchmark value considered by Allen and Donaldson (2022). We explore robustness of our results to this parameter choice in Appendix F.[30] Panel A of Table 2 displays the implied values of $\bar{\mu}^0$ and $\bar{\mu}^D$ given our baseline value of ζ.

[29] See Appendix E.1 for details on the estimation.

[30] We set the share of young-period consumption in lifetime utility, β, to 0.5. This parameter choice has no consequences for any of the outcomes in the quantitative analysis.

4.2.2 Second Step: Productivities, amenities, and fertility

Now, we calibrate the shape parameter of the initial productivity distribution, θ, and the experience premium parameter, A^o. In addition, we recover local amenities $u_{n,t}$, the complete path of scale parameters $\lambda_{n,s,t}$, and aggregate fertility f_t.

Consider the scale parameters $\lambda_{n,s,t}$. These terms summarize the knowledge stock in each location-sector and are at the core of the quantitative analysis: higher values of $\lambda_{n,s,t}$ indicate higher local income, greater ability to attract population, and greater potential for future innovation. While $\lambda_{n,s,t}$ are endogenous to the diffusion process, in this step we use the patent data to directly infer their paths over time. Later in the calibration, we will recover the exogenous shocks that fully rationalize those paths.

We begin by postulating a simple notion of patenting in the model that allows us to draw a connection between patent data and the path of $\lambda_{n,s,t}$. In particular, we assume that young individuals, when undergoing their imitation or innovation process outlined in Section 3.1.3, will file for a patent for their best innovation draw if it results in an idea of quality above a time-varying threshold, Q_t. In Appendix E.2, we show that, in this setting, the number of patents per young individual in (n, s) at time t is approximated by the following expression:

$$\frac{Pat_{n,s,t}}{L^y_{n,s,t}} \approx \Lambda_{n,s,t} Q_t^{-\theta}, \tag{27}$$

where $Pat_{n,s,t}$ denotes the number of patents filed at time t in (n, s), and $\Lambda_{n,s,t} \equiv \lambda_{n,s,t} - \lambda_{n,s,t-1}$. Given values for θ and Q_t, we can set the path of $\lambda_{n,s,t}$ and $u_{n,t}$ so to satisfy Equation (27) for all (n, s, t) and, simultaneously, match total population by location in all time periods. Equation (27) provides a simple mapping between the patent data and the path of $\lambda_{n,s,t}$, for which we provide extensive validation in Section 4.3 below.[31]

We set θ to match the weighted standard deviation of log-income per capita across cities in 1990, which is equal to 0.214. In Appendix Figure A.7, we show computationally that there is a unique value of θ that matches this data moment. The path of Q_t is calibrated to produce an annual growth in per capita income of 2%.[32] Lastly, we set $A^o = 2.11$, which implies a 1990 experience premium of 1.42 (Heathcote et al., 2010).[33]

[31] We choose units of the final good so that the geometric average of $\lambda_{n,s}$ is equal to one in the first time period.

[32] The term Q_t can be microfounded as a function of accumulated knowledge and the minimum inventive step an idea must satisfy to obtain a patent, as in Kortum (1997).

[33] The growth rate of $\lambda_{n,s,t}$ in the model is tightly linked to the rate of "creative destruction", defined as the rate at which old ideas are no longer used in production (Caballero and Jaffe, 1993):

$$\Delta_{n,s,t} \equiv \frac{\Lambda_{n,s,t}}{\lambda_{n,s,t-1} + \Lambda_{n,s,t}}.$$

In the model, a 2% average income growth per year corresponds to a yearly rate of "creative destruction" of about

We set the elasticity of local amenities to population, ω, to -0.15, following Eckert and Peters (2022). In Appendix F, we explore robustness of our results to this parameter choice. Given ω and the path of $u_{n,t}$, we can recover the exogenous component of local amenities, $\bar{u}_{n,t}$. Finally, we calibrate the fertility path, f_t, to match population totals by period. Panels B and C of Table 2 summarize values and targets of the parameters calibrated in this step.

4.2.3 Third Step: Gravity equation for knowledge flows

In the third step of the calibration, we recover the parameters controlling knowledge transmission costs, $d_{(m,r)\to(n,s),t}$. To this end, we derive a gravity representation for knowledge flows that we estimate using data on patent citations. We impose a flexible parametrization for knowledge diffusion frictions, accounting for geographical distance, a direct feedback from migration decisions into diffusion, and a full set of fixed effects for each pair of sectors of origin and destination, capturing a wide range of channels of diffusion and associated frictions.[34] We assume that these terms contribute multiplicatively to overall frictions:

$$d_{(m,r)\to(n,s),t} = e^{\delta^G \mathbf{1}_{m\neq n} + \delta^D D_{m,n} + \delta^M M_{m\to n,t} + \delta^K_{r\to s}}, \tag{28}$$

where $D_{m,n}$ denotes distance in 1,000 km between m and n, and $M_{m\to n,t}$ denotes n's migration exposure to m, defined as the share of old migrants into n that were born in m (and as equal to 0 for $m = n$). In Equation (28), δ^G controls the effectiveness of knowledge flows *across* locations relative to flows *within* locations, δ^D and δ^M control the sensitivity of transmission efficiency to geographical and migration proximity, respectively, and $\delta^K_{r\to s}$ controls the applicability of ideas from sector r for innovation in sector s.[35]

Combining Equations (11) and (28) and taking logs on both sides yields

$$\log(\eta_{(m,r)\to(n,s),t}) = \phi^0_{m,r,t} + \phi^1_{n,s,t} - \theta\delta^G \mathbf{1}_{m\neq n} - \theta\delta^D D_{m,n} - \theta\delta^M M_{m\to n,t} - \theta\delta^K_{r\to s}, \tag{29}$$

where ϕ^0 and ϕ^1 represent idea origin-time and idea destination-time fixed effects, respectively.

Equation (29) illustrates that bilateral citation probabilities $\eta_{(m,r)\to(n,s),t}$ depend on the composite parameters $\theta\delta^G$, $\theta\delta^D$, $\theta\delta^M$, and $\theta\delta^K_{r\to s}$. We estimate Equation (29) using data on patent citations across location-sector pairs in order to recover those composite parameters. To obtain

8.4%, which is close to the value of 7% estimated by Caballero and Jaffe (1993) for the mid-20th century.

[34] Results do not change significantly with richer frictions in which diffusion depends on the share of individuals in the destination location employed in the sector of origin of the idea.

[35] An alternative approach to calibrate frictions to diffusion is the one developed by Caballero and Jaffe (1993) and adapted by Cai et al. (2022a) to a setting with multiple countries and sectors. This approach delivers a separate parameter for each pair of country-sectors without relying on the parametric assumptions in Equation (28). However, implementing this approach requires observing citations at multiple lags, which is not feasible in our model, in which one period corresponds to 20 years.

a more precise measurement of the flow of knowledge among inventors, we only include citations added by applicants, limiting the sample to patents issued after 2000, when patent documents began reporting inventor-added citations separately from examiner-added citations (Alcacer and Gittelman, 2006). We compute $\eta_{(m,r)\to(n,s),t}$ as the proportion of citations from patents in (n,s) to patents in (m,r).[36]

We estimate this relationship using PPML and present the results in Appendix Table E.6. Using the estimated value of the composite parameters $\theta\delta^G$ and $\theta\delta^D$ in conjunction with the calibrated value of θ, we obtain $\delta^G = 1.43$ and $\delta^D = 0.037$. The coefficients indicate that knowledge flows are highly localized, with the effectiveness of transmission across locations at 1,000 km of distance, $e^{-\theta(\delta^G+\delta^D)}$, estimated to be approximately 0.14% of the effectiveness of transmission within the same location. However, the overall weight of ideas from outside locations may still be substantial in determining innovation in n, given that transmission can occur from *all* locations $m \neq n$. We also find a positive effect of migration exposure. A one percentage point increase in migration exposure is associated with 3.3% more citations to the origin commuting zone.

The same regression also yields a complete set of bilateral transmission costs across sectors ($\delta^K_{r\to s}$), as depicted in the heatmap of Appendix Figure E.14. These costs are lower within sectors (along the diagonal), but all pairs of sectors exhibit some degree of knowledge exchange that, in some cases, is far from negligible, such as for class-groups G1 ("Physics") and H1 ("Electricity"). Panel D of Table 2 summarizes the values of the parameters obtained at this step of the calibration.

4.2.4 Fourth Step: Technological waves and structural residuals

In the fourth calibration step, we combine the objects calibrated so far with the law of motion in Equation (9) to recover technological wave shocks $\alpha_{s,t}$ and structural residuals $\epsilon_{n,s,t}$. For all periods t, we first guess the full vector of technological wave shocks $\{\alpha_{s,t}\}_{s\in S}$. Given this guess, we use Equation (9) to recover the full set of structural residuals. This step rationalizes the path of $\lambda_{n,s,t}$ for any initial guess of $\{\alpha_{s,t}\}_{s\in S}$. For identification, we need to impose S additional conditions. Our working assumption is that in each period, variation in *average* productivity growth across sectors is explained by technological waves and their interaction with the endogenous process implied by Equation (9). Structural residuals, $\epsilon_{n,s,t}$, account for the remaining variation in productivity growth across locations for any given sector. Formally, this is captured by the following assumption:

[36]Note that the direction of the arrow from (m,r) to (n,s) denotes knowledge flows going from the *cited* patent to the *citing* patent. The total weight of each citing patent in our regression is one (i.e., the weight of each citation is proportional to the inverse of the total number of citations given by the citing patent). To account for the fact that *local* knowledge flows are more likely to be tacit and less likely to be captured by citations, we add an artificial citation to each patent's list of references to a local patent whose technology classes are identical to the citing patent. This assumption amounts to assuming that the weight of local tacit flows is inversely proportional to the number of inventor-provided citations. This also ensures that all patents, including those without backward citations, enter the estimation. Appendix E.3 provides additional details on the estimation.

Table 2: **Parameter values and targets**

Parameter	Value	Target / Source	Model	Data
Panel A: Estimated via gravity equation for migration flows				
$\bar{\mu}^0$	0.197	Table E.5		
$\bar{\mu}^D$	0.703	Table E.5		
Panel B: Externally calibrated parameters				
ζ	4.0	Allen and Donaldson (2022)		
ω	-0.15	Eckert and Peters (2022)		
Panel C: Internally calibrated parameters matching moments				
θ	4.48	s.d. log-income across CZs, 1990	0.214	0.214
A^o	2.11	Experience premium, 1990	1.42	1.42
Panel D: Estimated via gravity equation for knowledge flows				
δ^G	1.434	Table E.6		
δ^D	0.037	Table E.6		
δ^M	-0.745	Table E.6		
$\delta^K_{r\to s}$	Figure E.14	Table E.6		

Notes: Experience premium is taken from Heathcote et al. (2010). Standard deviation of log-income across CZs are author's calculations from the NHGIS.

Assumption A3. *The sets of adjusted structural residuals,* $\{\epsilon^{\theta}_{n,s,t}\}_{(n,s)\in N\times S, t\geq 0}$*, have a common average for each sector and time period that we normalize to one:*

$$\mathbb{E}\left[\epsilon^{\theta}_{\cdot,s,t}\right] = 1, \qquad \forall s \in S,\, t \geq 0. \tag{30}$$

In practice, our identification assumption implies that technological wave shocks are responsible for the common component of shifts in the knowledge stock in a given sector. Structural residuals account for the remainder of location-sector specific variation. Combining Equation (30) and the law of motion (9) enables us to recover a unique set of technological wave shocks, $\{\alpha_{s,t}\}_{s\in S}$ and structural residuals $\{\epsilon_{n,s,t}\}$.

The identification of technological waves is linked to the co-movement of knowledge stocks across locations within a given sector. Because we infer technological waves from patenting behavior, our model is mute regarding their potential causes. Several factors can drive technological waves, such as scientific advancements in specific fields or firms directing innovation efforts toward certain sectors based on their future demand forecasts (as suggested by Comin et al., 2019 when studying long-term patterns of sectoral productivity growth in the United States). Our identification strategy picks up as a technological wave any expected common sectoral shock that redirects innovation across industries.

Appendix Figure A.8 depicts the calibrated path of $\alpha_{s,t}$ for each technology class-group. In the early portion of the sample, "Agriculture" experiences the greatest negative shock, whereas "Chemistry" and "Electricity" exhibit visible growth. The latter portion of the sample reveals a general decline in fields associated with heavy manufacturing, such as "Transporting", "Engines", and "Chemistry", and significant expansions in "Health" and "Physics", with the latter containing the majority of IT and computing-related innovation.

4.3 Untargeted moments

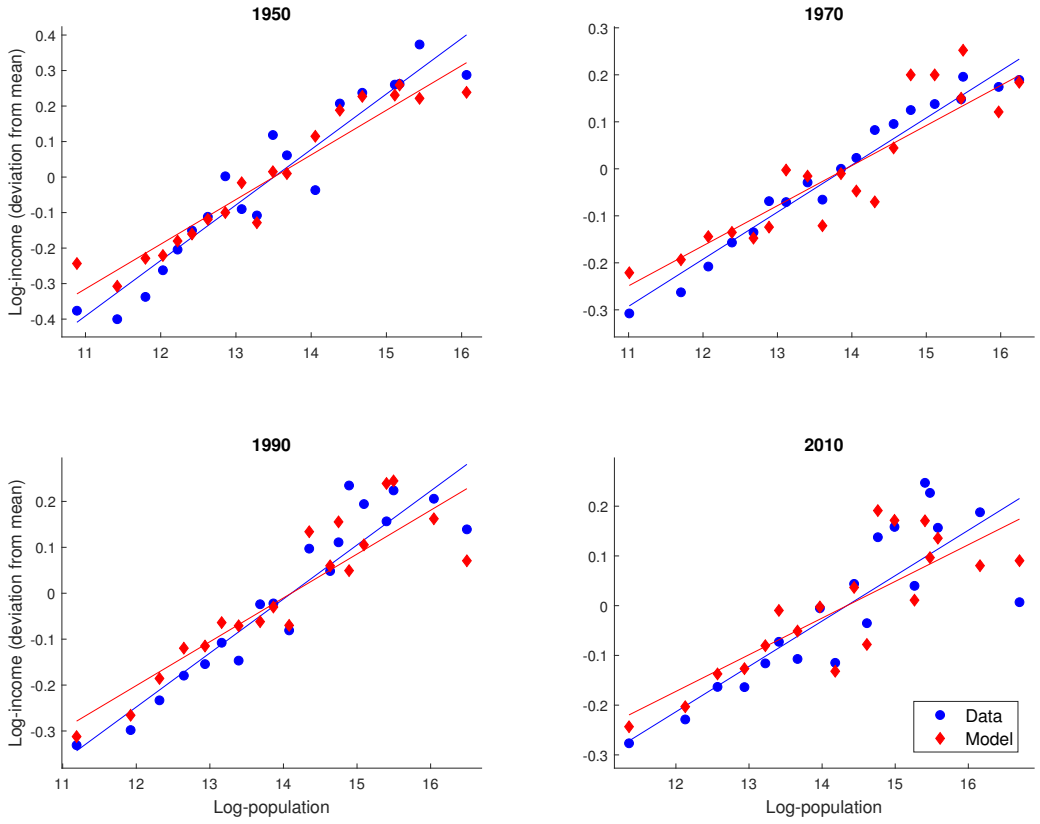
Some channels that are potentially relevant from a quantitative perspective, such as variation in local prices, are captured in our setting in reduced-form, allowing us to focus on our core mechanism of knowledge creation and diffusion. We now show that, despite its parsimony, our calibrated model is successful at accounting for some key moments that are not directly targeted. Since the calibration matches local population by construction, we focus our attention on the model's predictions on local income, which is the ultimate endogenous outcome of the process of innovation and diffusion, as well as the main driver of migration decisions.

The bin-scatter plots in Figure 3 depict the relationship between log-population and log-income in the model and in the data for all periods since 1950 (the earliest decade for which income data at the CZ level can be reliably constructed using the NHGIS, see Appendix B for details). The existence of a correlation between city size and productivity is a well-established empirical regularity (see, e.g., Glaeser and Gottlieb, 2009) that can result from a variety of theoretical mechanisms (e.g., sorting, variety, local productivity spillovers, and higher availability of inputs). While the model is silent on the mechanism underlying this correlation (other than the fact that more productive cities will *attract* more population), it is crucial for the model's quantitative performance that the elasticity of income with respect to population is empirically accurate. The figure shows that the model-implied and empirical elasticities are remarkably close, despite the fact that this moment is not targeted by the calibration.

Furthermore, the scatter plots in Appendix Figure A.9 display the relationship between log-income in the model and in the data across all cities between 1950 and 2010. Also in this case, the model-generated values are strongly predictive of their empirical counterparts, despite the latter not being directly used to inform the calibration. The R^2 of the underlying regressions range between 0.47 and 0.55.

While the model's ability to account for city-level variation in income is critical for its quantitative accuracy, the mechanism driving local growth ultimately relies on variation in income at the location-sector level. Local sectoral income is not directly observed throughout the sample. However, for the 1990 and 2010 decades, we can construct a reliable approximation for income at the location-sector level by combining IPUMS data on individual income with the crosswalk

Figure 3: **Population and income per capita: data and model (untargeted)**



Notes: Bin-scatter plots of the relationship between log-population and log-income per capita in the data (blue) and the model (red), weighted by population, in 1950 (top-left panel), 1970 (top-right panel), 1990 (bottom-left panel), and 2010 (bottom-right panel). Log-income is displayed as deviation from the mean.

between industry codes and IPC classes developed by Lybbert and Zolas (2014) (see Appendix B for details).[37] Appendix Figure A.10 shows bin-scatter plots of log-income in the model and in the data by location-sector, residualized with respect to location and sector fixed effects. The plots show a remarkable fit of this untargeted data, providing further validation of the mapping between patent data and local productivity, and our calibration strategy overall.

5 Quantitative results

In this section, we use the calibrated model to assess the role of technological waves and frictional idea diffusion on local growth. We then discuss two applications of our quantitative framework. First, we investigate the relationship between local specialization and growth volatility. Second,

[37] In the model decades before 1990, IPUMS does not provide consistent information either on income or geographical location (at a level that can be reliably mapped into 1990 CZs). For this reason, this validation can only be performed for 1990 and 2010.

we explore how future technological trends may reshape the economic geography of the United States in the coming decades.

5.1 Technological waves, local growth, and knowledge diffusion

Our primary objective is to quantify the contribution of frictions to idea diffusion to the empirical relationship between exposure to technological waves and local growth described in Section 2.3. To this end, we initialize the economy with the BGP conditions calibrated for the initial period (1890). We then compute the model forward until the final period (2010) under different assumptions for the evolution of the exogenous disturbances, and compare the resulting population dynamics.

First, we compute the model forward by feeding the full path of exogenous sectoral ($\alpha_{s,t}$) and local ($\epsilon_{n,s,t}$ and $\bar{u}_{n,t}$) shocks. This version of the model reproduces the data by construction. Thus, the resulting relationship between exposure to technological waves and local population growth is identical to the empirical one estimated in Section 2.3, as reported in column 1 of Table 3.

Second, we compute the model forward by feeding the full path of local shocks ($\epsilon_{n,s,t}$ and $\bar{u}_{n,t}$), while holding the aggregate technological wave shocks, $\alpha_{s,t}$, fixed at their 1890 BGP values. The population dynamics resulting from this exercise reflect those driven by factors orthogonal to technological waves, as captured by changes in the structural residuals and exogenous amenities. Then, we estimate an analogous regression to the empirical one (Equation 2), but with the counterfactual local growth derived from this exercise as the dependent variable. Notice that, in the absence of frictions to diffusion, the path of $\alpha_{s,t}$ would have no influence on the dynamics of local productivity and population. Hence, the difference between the empirical and counterfactual coefficients can be interpreted as the effect of technological wave shocks on local population growth via the endogenous mechanism of frictional knowledge diffusion.

The counterfactual coefficient, which is reported in column 2 of Table 3, is 65% lower than the empirical one. This suggests that the model's mechanism of frictional idea diffusion explains close to two-thirds of the empirical relationship between exposure to technological waves and local growth. The remaining correlation is explained by local shocks (structural errors and amenities) that are not part of the endogenous mechanism of the model.[38]

[38]The contribution of this mechanism throughout our sample can only be properly quantified by comparing a model with the full set of shocks to one in which technological waves are shut down. Agents' optimal innovation behavior implies a backward-looking law of motion for productivity: the productivity of a location-sector at time t is a function of all productivities at time $t-1$ and the set of technological wave shocks at time t (Equation 9). Hence, the model cannot endogenously account for major long-term shifts in the spatial distribution of economic activity caused by external factors, such as the World Wars and the Great Migration, which are captured by time-varying amenities ($\bar{u}_{n,t}$) and structural residuals ($\epsilon_{n,s,t}$). These terms, by reallocating people and changing the returns on innovation in the current period, influence our mechanism going forward. Comparing the model with the complete set of shocks to one in which technological waves are constant allows us to isolate the mechanism while maintaining the quantitative integrity of the experiment over the entire sample.

Table 3: **Population growth and technological wave shocks**

	Dependent var.: Growth rate of population under		
	Full model	Model without tech. waves	
			No diffusion across fields
	(1)	(2)	(3)
Exposure to tech. waves ($Exp_{n,t}$)	0.341***	0.119	0.237**
	(0.099)	(0.100)	(0.098)
Difference from empirical coefficient	-	0.222	0.104
Share explained by tech waves ($\alpha_{s,t}$)		65.1%	30.5%
Decomposition:			
- Share explained by tech. frictions			46.9%
- Share explained by geo. frictions			53.1%
Log-population (lags 1 and 2)	Yes	Yes	Yes
Other controls	Yes	Yes	Yes
Fixed effects	CD×T	CD×T	CD×T
# Obs.	2,852	2,852	2,852
R^2	0.54	0.54	0.54

Notes: CZ-level regression, 1910-2010, weighted by the share of population at the beginning of the period. Dependent variable defined as growth rate of population over 20 years. "CD×T" denotes Census Division-time fixed effects. Standard errors clustered at the CZ and the Census Division-time level in parenthesis. Exposure to the technological wave is defined as in Equation (1). Column 2 displays the counterfactual in which $\alpha_{s,t}$ are kept constant at their 1890 BGP values. Column 3 displays the counterfactual in which knowledge flows are restricted to within-field flows only. Controls include log-total patents, human capital, and industry composition. $^{**}p < 0.05$, $^{***}p < 0.01$.

5.1.1 Disentangling the effect of diffusion frictions across cities and knowledge fields

Our theory explains the impact of a city's exposure to technological waves on local population growth through two distinct channels. First, frictions to idea diffusion *across fields of knowledge* imply that sectors receiving favorable technological wave shocks will experience greater productivity growth. Consequently, cities with a greater concentration of expanding fields will experience greater productivity and population growth. This channel is emphasized by Equation (21), which is derived under the assumption that knowledge flows only occur within fields of knowledge, $\eta_{s\to(n,s),t} \approx 1$ (Assumption A2.1). Second, frictions to idea diffusion *across locations* imply that cities geographically close to a greater concentration of expanding fields will experience greater productivity growth in *all* sectors due to localized knowledge spillovers across fields.

To decompose the relative importance of these two channels, we re-calibrate technological wave shocks and structural residuals under the assumption that frictions across fields are sufficiently

strong that idea flows only occur *within* fields (e.g., $\delta^K_{r\to s} = \infty$ if $r \neq s$).[39] We then compute the model forward from the initial 1890 BGP until the last period (2010) under this assumption, keeping the aggregate technological wave shocks, $\alpha_{s,t}$, fixed at their BGP values. We finally estimate a regression of local growth on the exposure measure analogous to the empirical one.

Column 3 of Table 3 reports the corresponding estimates. The coefficient is 30.5% smaller than in the complete model of column 1. Compared to the estimate in column 2, this number suggests that approximately 47% of the overall mechanism can be attributed to frictions to diffusion across technological fields, while the remaining portion (approximately 53%) can be attributed to frictions to diffusion across locations.

5.1.2 Technological waves and the emergence of the current U.S. geography

We now use our framework to evaluate the historical contribution of technological waves in delineating the current economic geography of the United States.

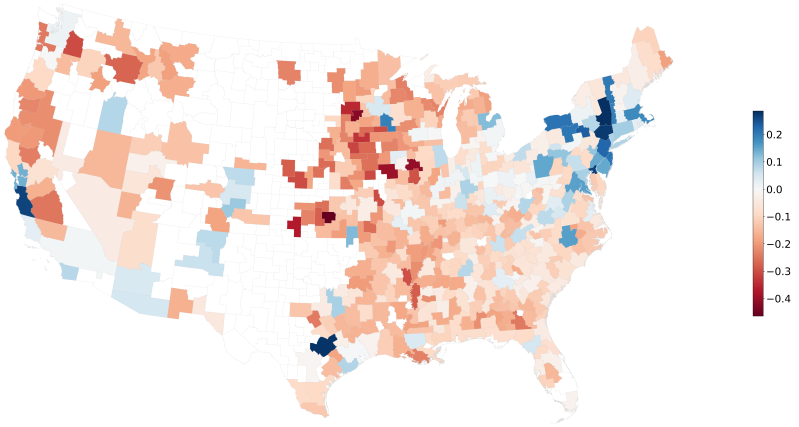
Figure 4 presents a map of the commuting zones in our sample colored in blue (red) if their realized growth rate between 1890 and 2010 is higher (lower) than the counterfactual growth rate obtained by keeping technological wave shocks, $\alpha_{s,t}$, fixed at their 1890 BGP values (as in column 2 of Table 3). Darker colors represent larger absolute differences between actual and counterfactual growth rates.

The map reveals that, throughout the last century, technological waves significantly bolstered the growth of cities in the Northeast, and promoted the rise of the most prominent modern innovation hubs, such as San Jose, Austin, and Raleigh-Durham. In the absence of these technological waves, the size of those commuting zones would have been 26.4%, 28.0%, and 15.7% smaller, respectively. Interestingly, Seattle—another major technological hub—does not appear to have been significantly affected. A plausible explanation is that the emergence of Seattle is mostly explained by idiosyncratic factors, such as the location choices of Microsoft and Amazon, which are largely orthogonal to our endogenous mechanism and are captured by the structural residual in the model.

The effect of technological waves throughout the whole sample period on cities in the Midwest and the Rust Belt is in general small in absolute value, reflecting the fact that technological waves have favored the growth of manufacturing-intensive sectors in the first half of the 20th century, and later reversed their fortunes. Cities in the Central States, with a higher concentration of agricultural activities, are the most negatively affected.

[39] Incidentally, this formulation is equivalent to assuming that diffusion across locations is *frictionless* (i.e., $\delta^G = \delta^D = \delta^M = 0$), since, in both cases, the reliance of local innovators on ideas from any given sector, $\eta_{r\to(n,s),t}$, will not depend on n and will hence be constant across cities. In both cases, this counterfactual isolates the role of technological frictions to diffusion from the role of geographical frictions.

Figure 4: **Contribution of technological waves, 1890-2010, to the current U.S. geography**



Notes: The map shows the difference between 2010 log-population in the data and in the counterfactual where $\alpha_{s,t}$ are kept at their 1890 GDP values. Blue (red) CZs indicate that realized population is higher (lower) than the counterfactual one.

5.2 Specialization and volatility of local growth

Our mechanism implies that growth in diversified cities is less volatile than in specialized ones. This happens for two reasons. First, due to frictions *across fields*, growth in a particular sector is primarily driven by shocks within that sector. Second, due to frictions *across locations*, the local reliance on ideas from a particular field is increasing in the local availability of ideas from that field. Diversified cities, whose composition is dispersed across many sectors, and whose innovators rely on ideas from a broader range of fields, experience less volatile productivity and ultimately population growth.

Exploring this link requires us to define a local specialization measure. Appendix G shows that under intuitive conditions, the variance of local population growth is proportional to the square of the Euclidean distance between the local and national vectors of innovators' reliance on ideas from each sector, which we use as our measure of specialization:

$$Spec_{n,t} \equiv \sum_{r \in S} \left(\eta_{r \to n,t} - \bar{\eta}_{r,t}\right)^2. \tag{31}$$

According to this measure, cities are perfectly diversified if the reliance of local innovators on ideas from each field r, denoted by $\eta_{r \to n,t} \equiv \sum_s \pi_{s|n,t} \eta_{r \to (n,s),t}$, is exactly equal to the economy-wise reliance of innovators on ideas from the same field, denoted by $\bar{\eta}_{r,t} \equiv \sum_{n,s} \pi_{n,s,t} \eta_{r \to (n,s),t}$.

To quantify the effect of specialization on local volatility, we run 5,000 simulations in which, starting from the last period of the model (2010), we randomly draw shocks to the growth rate of $\alpha_{s,t}$, $g^{\alpha}_{s,t}$, two periods into the future (i.e., until 2050). These shocks are drawn from a normal distribution with mean zero and standard deviation equal to the one of the calibrated $g^{\alpha}_{s,t}$ throughout

the sample. We then compute the counterfactual equilibrium for each simulation.

Column 1 of Appendix Table A.4 reports a regression of the standard deviation of local growth across the 5,000 simulations on the specialization measure in Equation (31). The volatility of population growth of specialized cities is significantly higher: the standard deviation for cities at the 90th percentile of the specialization distribution is 1.86 percentage points greater than for cities at the 10th percentile (the average volatility is 2.48 percentage points). To fix ideas, increasing specialization from the level of Boston (a highly diversified city) to the level of Austin (a highly specialized one) increases the standard deviation by 4.59 percentage points.

To disentangle the role of frictions across fields and across locations, we conduct the same experiment restricting knowledge diffusion to within-fields flows only (i.e., by imposing $\delta^K_{r\to s} = \infty$ if $r \neq s$). The results are reported in column 2 of Appendix Table A.4. The effect of specialization decreases by 24.6% compared to column 1. This implies that, while frictions across fields explain the majority of the effect of specialization on local volatility, frictions across locations, which dampen productivity growth fluctuations in more diversified cities, account for a non-negligible portion of the effect.

5.3 Impact of future changes in frictions to diffusion and technological waves

Finally, we use our model to examine how the U.S. economic geography may change in the coming decades in response to possible future technological transformations. Specifically, we project local growth until 2050 under various hypotheses on the evolution of frictions to diffusion ($d_{(m,r)\to(n,s),t}$) and technological possibilities in different sectors ($\alpha_{s,t}$), and compare the results to a baseline scenario in which the importance of all sectors is held constant at their 2010 levels.

5.3.1 Declining frictions to diffusion

We begin by considering a scenario in which geographical frictions to diffusion drop significantly, making knowledge transmission across locations more effective. While ICT was already pervasive in the latest years of our sample, the decline in long-distance communication costs saw a marked acceleration in the early 2020s, driven by a proliferation of online communication tools during the COVID-19 pandemic, and is expected to continue in the coming years (Barrero et al., 2021; Yang et al., 2022). We examine the effect of this trend by simulating a 50% drop in adjusted diffusion frictions across locations ($(d_{(m,r)\to(n,s),t})^\theta$, for all $m \neq n$).[40] Panel (a) of Appendix Figure A.11 illustrates the results. Compared to the baseline, CZs in blue (red) experience a net gain

[40] In the experiments described in this section, while the magnitudes of the shocks is set arbitrarily, the resulting qualitative patterns do not depend on those particular choices.

(loss) of population between 2010 and 2050. Major cities, including most of the main innovation hubs (such as Austin, San Jose, and Raleigh-Durham) experience relative losses in population, as the drop in diffusion frictions erodes their location-specific advantage in innovation. The drop in communication costs results in a convergence of population growth towards less dense locations in the Central States and the West. This prediction is in line with recent empirical evidence on the effect of remote work on the spatial concentration of economic activity (see, e.g., Ramani and Bloom, 2021 and Delventhal and Parkhomenko, 2023).

We next consider a scenario in which recent advances in Artificial Intelligence result in improved possibilities of recombination of ideas across different technological fields. This trend is already visible in several areas of innovation, most notably in the life-sciences, and is expected to become more pervasive in the coming decades (Bianchini et al., 2022; Agrawal et al., 2024). We simulate this scenario by assuming a 50% drop in adjusted diffusion frictions across sectors ($(d_{(m,r)\to(n,s),t})^{\theta}$, for all $r \neq s$), and display the results in panel (b) of Appendix Figure A.11. Surprisingly, the effect is similar to the one in the first scenario, with larger cities and major innovation hubs experiencing a relative loss in population to the benefit of less dense regions. The specialization of fast-growing cities in expanding fields prevents them from leveraging these new opportunities of recombination, which erodes their advantage compared to less specialized cities.

5.3.2 Future technological waves

As a last step, we examine the geographical effects of three plausible scenarios of technological trends involving specific fields.

In the first scenario, we assume that class-group B2 ("Transporting") experiences a positive technological wave shock that raises $\alpha_{s,t}$ by 20% compared to its 2010 value.[41] This scenario is conceivable if new advances in transit technologies and autonomous vehicles cause transportation innovation to resume its central role. Panel (a) of Appendix Figure A.12 shows the results. Rust Belt cities are the areas best positioned to benefit from this transformation. Detroit's population would increase by 11.5% relative to the baseline, and other centers of manufacturing would also benefit to a lesser extent. A relative loss of population would be experienced by the three knowledge hubs of Austin (-5.6%), San Jose (-4.1%), and Seattle (-1.3%).

An alternative way of modeling this scenario is to assume that ideas from "Transporting" become more relevant for innovation in either G1 ("Physics") or H1 ("Electricity") and vice versa, due to the gradual incorporation of IT components in electric and autonomous vehicles. We model this by assuming a 50% decrease in adjusted diffusion frictions ($(d_{(m,r)\to(n,s),t})^{\theta}$) from (to) B2 to (from) both G1 and H1. In this case, we maintain the 2010 value of sectoral importance ($\alpha_{s,t}$). The results are displayed in panel (b) of Appendix Figure A.12. While the magnitude of the overall

[41] A 20% shock corresponds to approximately 3.5 times the standard deviation of calibrated shocks.

effects is smaller, also in this scenario Detroit experiences a population increase (+0.6%), while San Jose experiences a net population loss (-0.1%). San Jose has become increasingly specialized, preventing it from capitalizing on cross-field spillovers, whereas Detroit's economy has recently diversified toward fields G1 and H1 and is thus exposed to ideas from those fields.

In the second scenario, we simulate a large (+20%) positive technological wave shock to class-group A3 ("Health; Life-Saving; Amusement," which includes innovation related to pharmaceuticals and medical sciences) possibly in response to new global health challenges. The results are depicted in panel (a) of Appendix Figure A.13. Major CZs in the Northeast, such as Boston (+5.8%) and Providence (+16.9%), and in California, such as Los Angeles (+2.8%) and San Francisco-Oakland (+0.6%), would experience a net inflow of population at the expense of IT clusters such as Austin (-8.9%), San Jose (-6.5%), and Seattle (-3.5%).

In the third scenario, we assume that class-group A1 ("Agriculture") regains centrality by experiencing a 20% technological wave shock. This scenario may result from tightening regulatory constraints and shifting demand toward sustainable farming, possibly in response to global challenges such as climate change. Results are displayed in panel (b) of Appendix Figure A.13. Under this scenario, population growth shifts from the East and West coasts and Rust Belt to the Central States (Des Moines has the greatest net gain among the major CZs) resulting in an overall convergence across regions.

6 Conclusions

The economic geography of countries is characterized by rich and heterogeneous dynamics. Some cities remain prosperous over long periods, whereas others experience episodes of rapid expansion and contraction. Understanding the drivers of these rich dynamics of local growth is of primary importance for policy and welfare. In this paper, we explore the hypothesis that these patterns are in part driven by cities' distribution of knowledge across fields, frictions to idea diffusion, and a continuously evolving technological landscape.

We provide reduced-form evidence of a positive correlation between local growth and exposure to technological waves in the U.S. during the twentieth century. To explain this relationship, we develop a tractable framework that combines elements from quantitative spatial models and theories of endogenous growth via innovation and idea diffusion. The quantitative results suggest that our proposed mechanism of frictional knowledge diffusion can account for close to two-thirds of the reduced-form correlation between local population growth and exposure to technological waves. Counterfactual experiments suggest that future technological scenarios may have large and heterogeneous geographical effects. These results imply that the process of frictional knowledge diffusion plays an important role in explaining why the technological landscape and the economic

geography appear to be deeply intertwined.

The framework developed in this paper opens up several avenues for future research. First, our quantitative results suggest that residual factors contribute significantly to the dynamics of local innovation and to the variation in city growth. Unpacking the residual term, by accounting for endogenous innovation effort or investment, firms' location choice, and policy response to local shocks, is a promising next step. Second, our framework can be easily extended to allow for ex-ante heterogeneous agents. Since the response to technological wave shocks is likely to vary across demographic groups, an extended model can be used to study the implications of our mechanism for the distribution of income and welfare.

Our results can also be used to inform the design of local and place-based interventions. Policy efforts aimed at shaping the local sectoral composition have become increasingly common in recent years (Slattery and Zidar, 2020). Our analysis suggests that these policies, by affecting the local availability of ideas from different knowledge fields, influence the sensitivity of local growth to future changes in the technological environment. The model highlights two policy tradeoffs. First, there is an inherent tension between short-term growth, which can be fostered by specializing in the sectors that currently offer the best innovation opportunities, and long-term stability, which can be achieved by diversifying the local portfolio of activities. Second, the aggregate benefits of increasing the concentration of innovation conceal the fact that these benefits are unevenly distributed in space. These tradeoffs suggest that an optimal growth policy critically depends on the planner's time horizon and geographical scope, and underscore a potentially important role for insurance and redistributive mechanisms in the design of local and national development policies.

References

Adao, R., C. Arkolakis, and F. Esposito (2020): "General Equilibrium Effects in Space: Theory and Measurement," Tech. rep., Department of Economics, Tufts University.

Agrawal, A., J. McHale, and A. Oettl (2024): "Artificial intelligence and scientific discovery: A model of prioritized search," *Research Policy*, 53, 104989.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015): "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 83, 2127–2189.

Akcigit, U. and W. R. Kerr (2018): "Growth through heterogeneous innovations," *Journal of Political Economy*, 126, 1374–1443.

Alcacer, J. and M. Gittelman (2006): "Patent citations as a measure of knowledge flows: The influence of examiner citations," *The review of economics and statistics*, 88, 774–779.

Allen, T. and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129, 1085–1140.

Allen, T. and D. Donaldson (2022): "Persistence and path dependence in the spatial economy," Tech. rep.

Andrews, M. J. (2021): "Historical patent data: A practitioner's guide," *Journal of Economics & Management Strategy*, 30, 368–397.

Arkolakis, C., S. K. Lee, and M. Peters (2020): "European immigrants and the United States' rise to the technological frontier," *Working paper*.

Atkeson, A. and A. Burstein (2019): "Aggregate implications of innovation policy," *Journal of Political Economy*, 127, 2625–2683.

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

Balland, P.-A., D. Rigby, and R. Boschma (2015): "The technological resilience of US cities," *Cambridge Journal of Regions, Economy and Society*, 8, 167–184.

Barrero, J. M., N. Bloom, and S. J. Davis (2021): "Why working from home will stick," Tech. rep., National Bureau of Economic Research.

Bartik, T. J. (1991): *Who benefits from state and local economic development policies?*, WE Upjohn Institute for Employment Research.

Baum-Snow, N. (2023): "Constraints on city and neighborhood growth: The central role of housing supply," *Journal of Economic Perspectives*, 37, 53–74.

Berkes, E. (2018): "Comprehensive Universe of U.S. Patents (CUSP): Data and Facts," mimeo, available at: https://sites.google.com/view/enricoberkes/work-in-progress.

Berkes, E. and R. Gaetani (2023): "Income segregation and the rise of the knowledge economy," *American Economic Journal: Applied Economics*, 15, 69–102.

Bianchini, S., M. Müller, and P. Pelletier (2022): "Artificial intelligence in science: An emerging general method of invention," *Research Policy*, 51, 104604.

Bleakley, H. and J. Lin (2012): "Portage and path dependence," *The quarterly journal of economics*, 127, 587–644.

Borusyak, K., R. Dix-Carneiro, and B. Kovak (2022): "Understanding Migration Responses to Local Shocks," *Available at SSRN 4086847*.

Bostic, R. W., J. S. Gans, and S. Stern (1997): "Urban productivity and factor growth in the late nineteenth century," *Journal of urban economics*, 41, 38–55.

Brezis, E. S. and P. R. Krugman (1997): "Technology and the life cycle of cities," *Journal of Economic Growth*, 2, 369–383.

Buera, F. J. and R. E. Lucas (2018): "Idea flows and economic growth," *Annual Review of Economics*, 10, 315–345.

Buera, F. J. and E. Oberfield (2020): "The global diffusion of ideas," *Econometrica*, 88, 83–114.

Burchardi, K. B., T. Chaney, T. A. Hassan, L. Tarquinio, and S. J. Terry (2020): "Immigration, innovation, and growth," Tech. rep., National Bureau of Economic Research.

Caballero, R. J. and A. B. Jaffe (1993): "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," in *NBER Macroeconomics Annual 1993, Volume 8*, National Bureau of Economic Research, Inc, NBER Chapters, 15–86.

Cai, J. and N. Li (2019): "Growth through inter-sectoral knowledge linkages," *The Review of Economic Studies*, 86, 1827–1866.

Cai, J., N. Li, and A. M. Santacreu (2022a): "Knowledge diffusion, trade, and innovation across countries and sectors," *American Economic Journal: Macroeconomics*, 14, 104–145.

Cai, S., L. Caliendo, F. Parro, and W. Xiang (2022b): "Mechanics of Spatial Growth," Tech. rep., National Bureau of Economic Research.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018): "The impact of regional and sectoral productivity changes on the US economy," *The Review of economic studies*, 85, 2042–2096.

Catalini, C., C. Fons-Rosen, and P. Gaulé (2020): "How do travel costs shape collaboration?" *Management Science*, 66, 3340–3360.

Comin, D., D. Lashkari, and M. Mestieri (2019): "Structural Change in Innovation," Tech. rep.

Davis, D. R. and J. I. Dingel (2019): "A spatial knowledge economy," *American Economic Review*, 109, 153–70.

Davis, D. R. and D. E. Weinstein (2002): "Bones, bombs, and break points: the geography of economic activity," *American Economic Review*, 92, 1269–1289.

De la Croix, D., M. Doepke, and J. Mokyr (2018): "Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy," *The Quarterly Journal of Economics*, 133, 1–70.

Delventhal, M. and A. Parkhomenko (2023): "Spatial implications of telecommuting," *Available at SSRN 3746555*.

Desmet, K., R. E. Kopp, S. A. Kulp, D. K. Nagy, M. Oppenheimer, E. Rossi-Hansberg, and B. H. Strauss (2018a): "Evaluating the economic cost of coastal flooding," Tech. rep., National Bureau of Economic Research.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018b): "The geography of development," *Journal of Political Economy*, 126, 903–983.

Desmet, K. and J. Rappaport (2017): "The settlement of the United States, 1800–2000: the long transition towards Gibrat's law," *Journal of Urban Economics*, 98, 50–68.

Desmet, K. and E. Rossi-Hansberg (2014): "Spatial development," *American Economic Review*, 104, 1211–43.

Diamond, R. (2016): "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000," *American Economic Review*, 106, 479–524.

Duranton, G. (2007): "Urban evolutions: The fast, the slow, and the still," *American Economic Review*, 97, 197–221.

Duranton, G. and D. Puga (2001): "Nursery cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review*, 91, 1454–1477.

Eckert, F. and M. Peters (2022): "Spatial structural change," Tech. rep., National Bureau of Economic Research.

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

Giannone, E. (2022): "Skilled-biased technical change and regional convergence," *Working paper*.

Glaeser, E. L. (2005): "Reinventing Boston: 1630–2003," *Journal of Economic Geography*, 5, 119–153.

Glaeser, E. L. and J. D. Gottlieb (2009): "The wealth of cities: Agglomeration economies and spatial equilibrium in the United States," *Journal of economic literature*, 47, 983–1028.

Glaeser, E. L. and J. Gyourko (2005): "Urban decline and durable housing," *Journal of political economy*, 113, 345–375.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer (1992): "Growth in cities," *Journal of political economy*, 100, 1126–1152.

Glaeser, E. L. and A. Saiz (2003): "The rise of the skilled city," .

Greenstone, M., R. Hornbeck, and E. Moretti (2010): "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," *Journal of Political Economy*, 118, 536–598.

Hanlon, W. W. and S. Heblich (2022): "History and urban economics," *Regional Science and Urban Economics*, 94, 103751.

Hayes, D. P. (1992): "The growing inaccessibility of science," *Nature*, 356, 739–740.

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006," *Review of Economic dynamics*, 13, 15–51.

Heblich, S., S. J. Redding, and D. M. Sturm (2020): "The making of the modern metropolis: evidence from London," *The Quarterly Journal of Economics*, 135, 2059–2133.

Holmes, T. J. and J. J. Stevens (2004): "Spatial distribution of economic activities in North America," in *Handbook of regional and urban economics*, Elsevier, vol. 4, 2797–2843.

Hornbeck, R. and E. Moretti (2018): "Who benefits from productivity growth? Direct and indirect effects of local TFP growth on wages, rents, and inequality," Tech. rep., National Bureau of Economic Research.

Huang, J. and Y. Zenou (2020): "Key Sectors in Endogeneous Growth," .

Jacobs, J. (1969): *The economy of cities*, Vintage international, Random House.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993): "Geographic localization of knowledge spillovers as evidenced by patent citations," *the Quarterly journal of Economics*, 108, 577–598.

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2021): "Measuring Technological Innovation over the Long Run," *American Economic Review: Insights*, 3, 303–20.

Kerr, W. R. and S. D. Kominers (2015): "Agglomerative forces and cluster shapes," *Review of Economics and Statistics*, 97, 877–899.

Kleinman, B., E. Liu, and S. J. Redding (2021): "Suficient Statistics for Dynamic Spatial Economics," .

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

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017): "Technological Innovation, Resource Allocation, and Growth*," *The Quarterly Journal of Economics*, 132, 665–712.

Kortum, S. S. (1997): "Research, Patenting, and Technological Change," *Econometrica*, 65, 1389–1420.

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

——— (2023): "Global innovation and knowledge diffusion," *American Economic Review: Insights*, 5, 494–510.

Lucas, R. E. and B. Moll (2014): "Knowledge growth and the allocation of time," *Journal of Political Economy*, 122, 1–51.

Lybbert, T. J. and N. J. Zolas (2014): "Getting patents and economic data to speak to each other: An 'algorithmic links with probabilities' approach for joint analyses of patenting and economic activity," *Research Policy*, 43, 530–542.

Manson, S., J. Schroeder, D. Van Riper, T. Kugler, and S. Ruggles (2021): "IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS," http://doi.org/10.18128/D050.V16.0.

Marshall, A. (1890): *The Principles of Economics*, McMaster University Archive for the History of Economic Thought.

Michaels, G., F. Rauch, and S. J. Redding (2012): "Urbanization and structural transformation," *The Quarterly Journal of Economics*, 127, 535–586.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018): "Commuting, migration, and local employment elasticities," *American Economic Review*, 108, 3855–90.

Moretti, E. (2012): *The new geography of jobs*, Houghton Mifflin Harcourt.

Morris-Levenson, J. and M. Prato (2021): "The Origins of Regional Specialization," Tech. rep.

Murata, Y., R. Nakajima, R. Okamoto, and R. Tamura (2014): "Localized Knowledge Spillovers and Patent Citations: A Distance-Based Approach," *Review of Economics and Statistics*, 96, 967–985.

Nagy, D. K. (2017): "City Location and Economic Development," mimeo, available at: https://sites.google.com/site/davidknagy/research.

Owens III, R., E. Rossi-Hansberg, and P.-D. Sarte (2020): "Rethinking detroit," *American Economic Journal: Economic Policy*, 12, 258–305.

Pauly, S. and F. Stipanicic (2022): "The creation and diffusion of knowledge: Evidence from the Jet Age," .

Perla, J. and C. Tonetti (2014): "Equilibrium imitation and growth," *Journal of Political Economy*, 122, 52–76.

Ramani, A. and N. Bloom (2021): "The Donut Effect of Covid-19 on Cities," *NBER Working Paper*.

Redding, S. J. and E. Rossi-Hansberg (2017): "Quantitative spatial economics," *Annual Review of Economics*, 9, 21–58.

Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek (2021): "IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021," https://doi.org/10.18128/D010.V11.0.

Schubert, G. (2021): "House price contagion and us city migration networks," *Working paper*.

Simon, C. J. and C. Nardinelli (2002): "Human capital and the rise of American cities, 1900–1990," *Regional Science and Urban Economics*, 32, 59–96.

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

Voigtländer, N. and H.-J. Voth (2013): "The three horsemen of riches: Plague, war, and urbanization in early modern Europe," *Review of Economic Studies*, 80, 774–811.

Yang, L., D. Holtz, S. Jaffe, S. Suri, S. Sinha, J. Weston, C. Joyce, N. Shah, K. Sherman, B. Hecht, et al. (2022): "The effects of remote work on collaboration among information workers," *Nature human behaviour*, 6, 43–54.

Online Appendix
(for online publication only)

Technological Waves, Knowledge Diffusion, and Local Growth

by Enrico Berkes, Ruben Gaetani, and Martí Mestieri

September 2024

A Additional tables and figures

Table A.1: **IPC Class-Groups**

Class ID	Class Group	IPC Class Range	Label
1	A1	A01-A24	Agriculture - Foodstuffs; Tobacco
2	A2	A41-A47	Personal or Domestic Articles
3	A3	A61-A99	Health; Life-Saving; Amusement
4	B1	B01-B44	Separating; Mixing - Shaping - Printing
5	B2	B60-B68	Transporting
-	B3	B81-B99	Microstructural Technology; Nanotechnology
6	C1	C01–C30	Chemistry - Metallurgy
-	C2	C40-C99	Combinatorial Technology
-	D1	D01-D07	Textiles - Paper
7	E1	E01-E99	Building - Earth or Rock Drilling; Mining
8	F1	F01-F17	Engines or Pumps - Engineering in General
9	F2	F21-F99	Lighting; Heating - Weapons; Blasting
10	G1	G01-G16	Physics
-	G2	G21-G99	Nuclear Physics; Nuclear Engineering
11	H1	H01-H99	Electricity

Notes: This table provides label and a mapping to the original IPC classes for the class-groups used for the empirical and quantitative analysis of this paper. Groups B3, C2, D1, and G2 are excluded from the sample since they are either negligible in size or they cover innovation in fields, such as nuclear physics, was acquired only in the later portion of the sample.

Table A.2: **Summary Statistics**

Variable	Obs.	Mean	Std. Dev.	Min	Max
Population	3,880	295,188.9	807,644.2	12	1.79e+07
Log-population	3,880	11.63	1.32	2.48	16.70
Population growth	2,910	.225	.311	-.684	2.428
Total patents	3,395	1,104.56	4,740.96	0	77,956
Exposure to tech. waves	2,910	-.107	.089	-.443	.405
Industrial composition	2,882	-.150	.157	-.681	.254

Notes: Summary statistics refer to the period 1870-2010 (Population, log-population), 1890-2010 (total patents), or 1910-2010 (remaining variables).

Table A.3: **Technological waves and city growth: Alternative specifications**

	Growth rate of population			
	(1)	(2)	(3)	(4)
Exposure to tech. waves, $Exp_{n,t}$	0.341***	0.387***	0.306***	0.258***
	(0.099)	(0.090)	(0.080)	(0.073)
Log-total patents	0.006	-0.002	0.017	-0.011
	(0.014)	(0.015)	(0.013)	(0.013)
Industry composition	0.573***	0.477***	0.698***	0.385***
	(0.113)	(0.098)	(0.120)	(0.102)
Human capital (ranking)	0.020	0.212***	0.080	0.289***
	(0.066)	(0.049)	(0.063)	(0.043)
Log-population (lags 1 and 2)	Yes	Yes	Yes	Yes
Fixed effects	CD×T	CD×T, CZ	CD×T	CD×T, CZ
Weighted	Yes	Yes	No	No
# Obs.	2,852	2,852	2,852	2,852
R^2	0.544	0.775	0.446	0.716

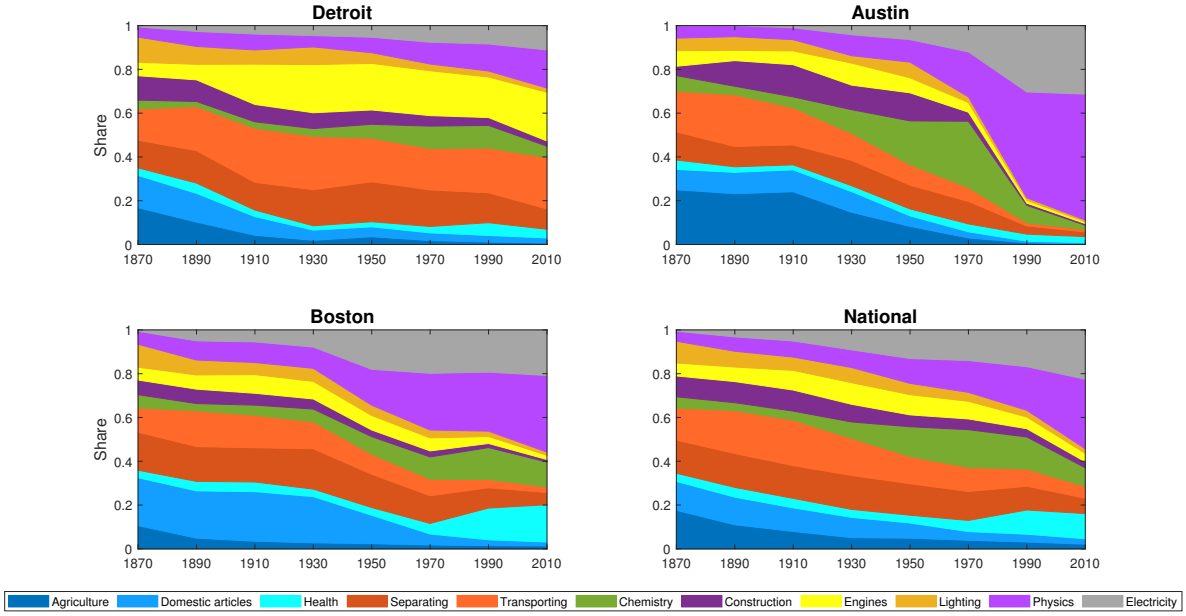
Notes: CZ-time-level regression, 1910-2010, weighted by share of population at the beginning of the period (columns 1 and 2) and unweighted (columns 3 and 4). Dependent variable defined as growth rate of population over 20 years. "CD×T" denotes Census Division-time fixed effects, "CZ" denotes CZ fixed effects. Standard errors clustered at the CZ and the Census Division-time level in parenthesis. ***$p < 0.01$.

Table A.4: **Specialization and volatility of population growth**

	Dependent var.: Standard deviation across simulations	
	(1)	(2)
Specialization in 2010 ($Spec_{n,2010}$)	0.841***	0.634***
	(0.043)	(0.045)
Log-population in 2010	-0.001**	-0.003***
	(0.0004)	(0.0004)
Diffusion across fields	Yes	No
$Spec_{n,2010}$, 90th-10th perc.	0.0221	0.0221
Mean of dependent var.	0.0248	0.0241
Fixed effects	CD	CD
# Obs.	485	485
R^2	0.53	0.47

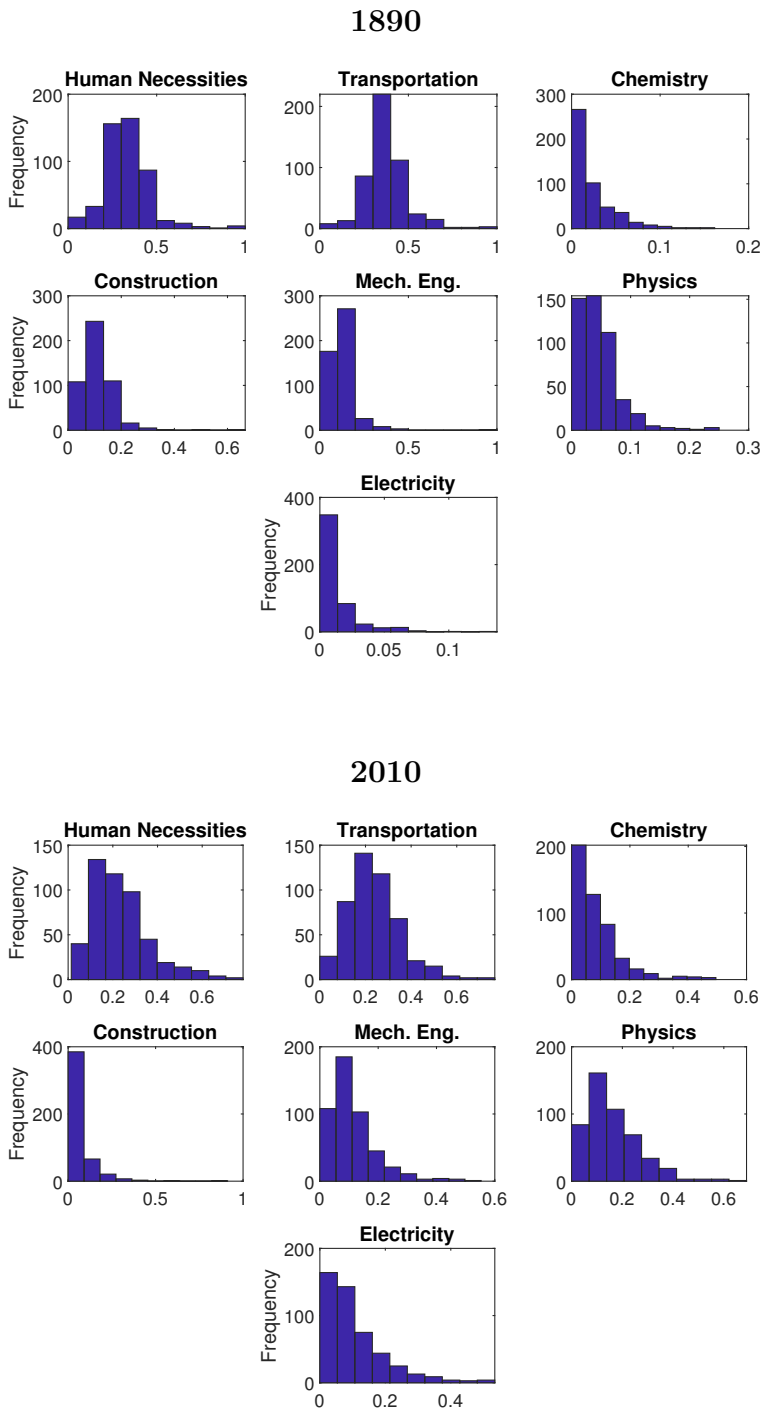
Notes: OLS estimates, weighted by population in 2010. Specialization is defined as in Equation (31). The dependent variable is defined as the city-level standard deviation of population growth across 5,000 simulations. "CD" denotes Census Division fixed effects. $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure A.1: **Composition of the technological output (class-groups)**



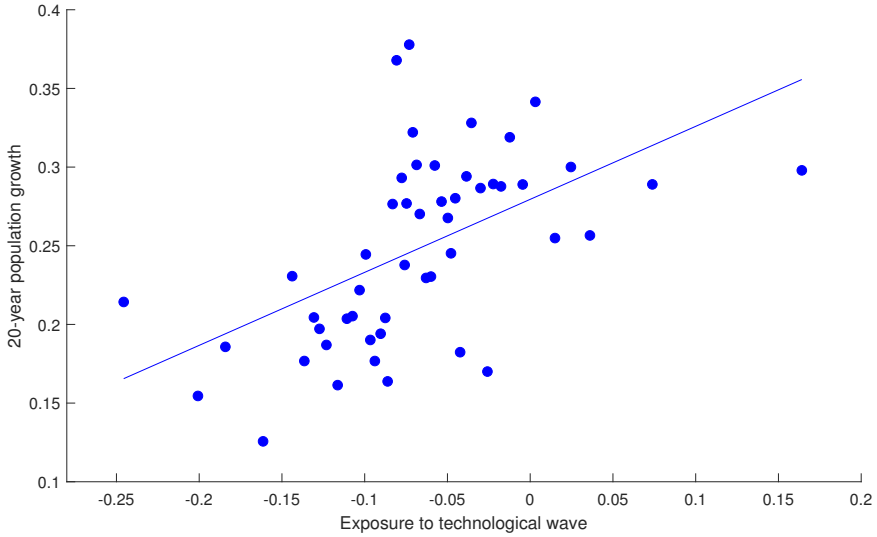
Notes: Composition of patenting output across the 11 IPC class-groups in Appendix Table A.1. Patent count for year t is constructed as the sum of patents filed between $t-10$ and $t+9$. Class names are abbreviated. The full description of each class is available at https://www.wipo.int/classifications/ipc/en/.

Figure A.2: **Distribution of patent shares across cities in 1890 and 2010**



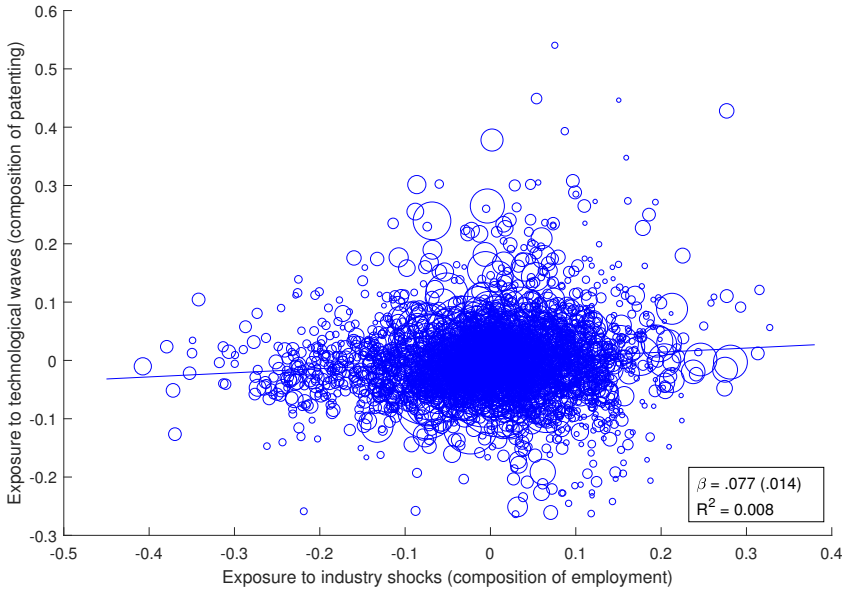
Notes: Distribution of patenting shares across cities for the seven main IPC classes in 1890 (top panel) and 2010 (bottom panel). Patent count for year t is constructed as the sum of patents filed between $t-10$ and $t+9$. Class names are abbreviated. The full description for each class is available at https://www.wipo.int/classifications/ipc/en/.

Figure A.3: **Technological waves and city growth**



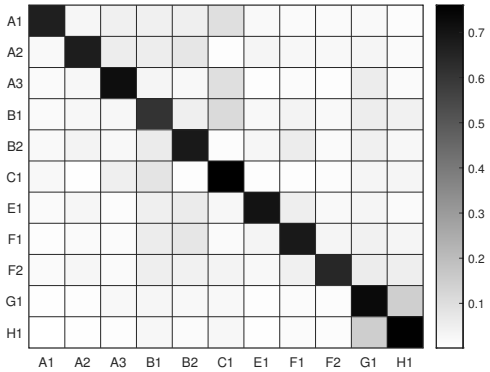
Notes: Bin-scatter plot of exposure to the technological wave, as defined in Equation (1), and 20-year population growth, 1910-2010, weighted by share of population at the beginning of the period. The bin-scatter plot is residualized with respect to Census Division-time fixed effects and two lags of log-population.

Figure A.4: **Exposure to technological waves and local industry composition**



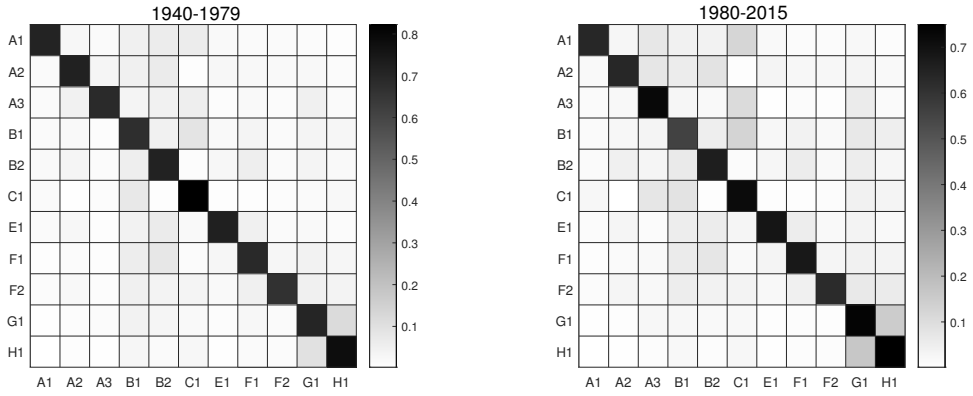
Notes: Scatter plot of exposure to the technological wave, as defined in Equation (1), and exposure to industry shocks, constructed using employment by industry, using the 12 main industries in the 1950 Census Bureau industrial classification system, weighted by share of population at the beginning of the period. Both variables are residualized with respect to two lags of log-population and Census Division-time fixed effects.

Figure A.5: **Patent citations across fields**



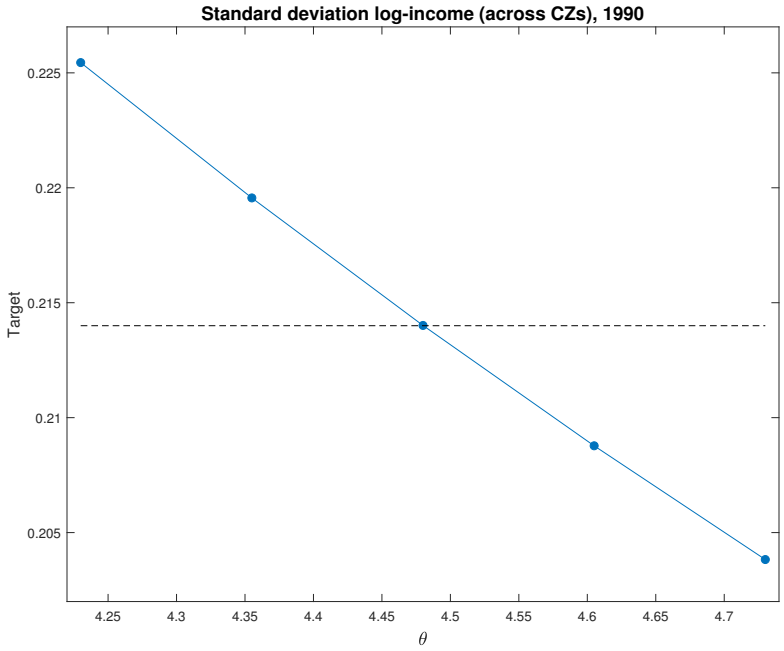
Notes: Probability that patents from the class-group on the vertical axis cite patents from the class-group on the horizontal axis. Probabilities are computed using patents filed since 1940 in which at least one inventor is in one of the 485 commuting zones in the main sample. Each citation is weighted by the inverse of the total number of citations given by the citing patent. Class-groups are described in Appendix Table A.1.

Figure A.6: **Patent citations across fields: Early vs late samples**



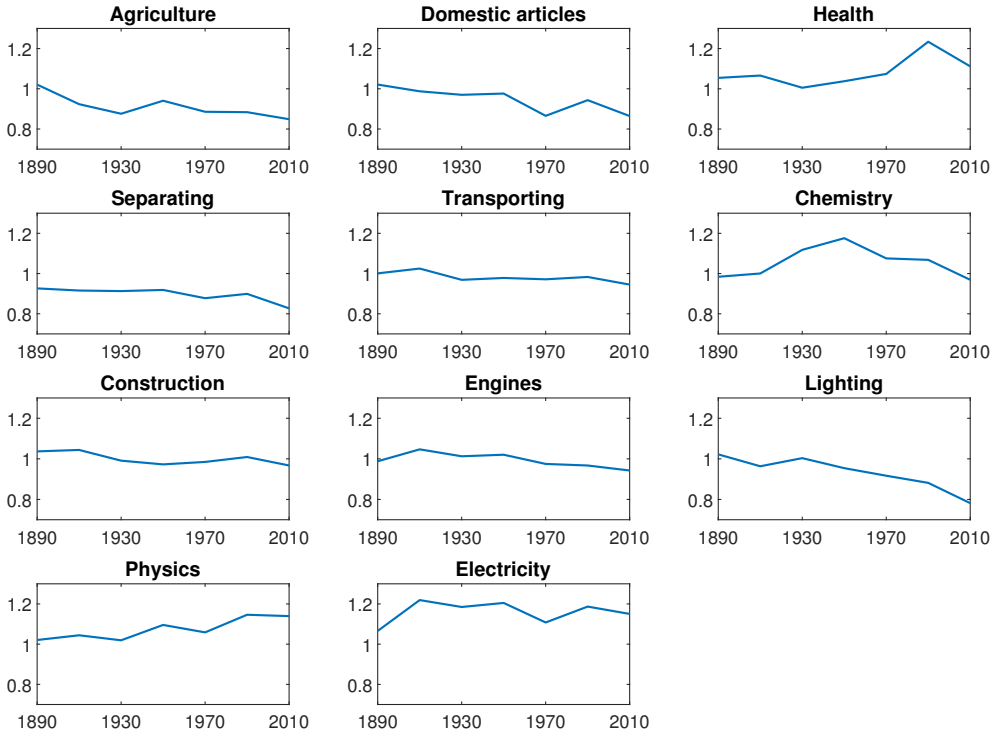
Notes: Probability that patents from the class-group on the vertical axis cite patents from the class-group on the horizontal axis. Probabilities are computed using patents filed since 1940 in which at least one inventor is in one of the 485 commuting zones in the main sample. Each citing patent is assigned a total weight of one. Class groups are described in Appendix Table A.1.

Figure A.7: **Identification of θ**



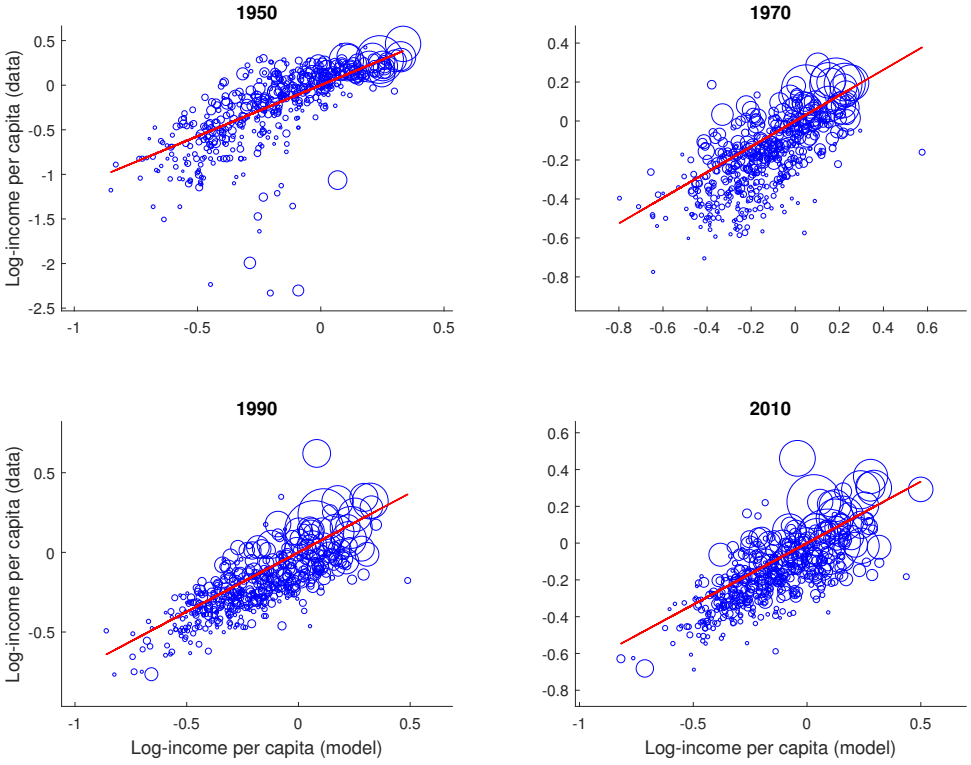
Notes: Standard deviation of log income across CZs in the data (horizontal dotted line) and in the model (blue marked line) for different values of θ.

Figure A.8: **Calibrated technological wave shocks ($\alpha_{s,t}$)**



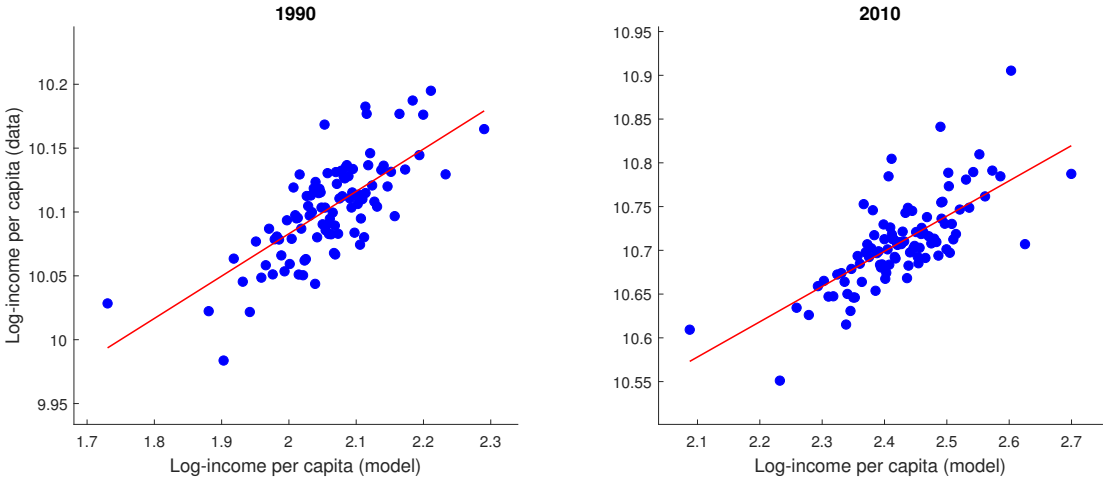
Notes: Paths of calibrated technological wave shocks ($\alpha_{s,t}$), 1890-2010. Class names are abbreviated. The full description of each class is available at https://www.wipo.int/classifications/ipc/en/.

Figure A.9: **Income per capita: data and model (untargeted)**



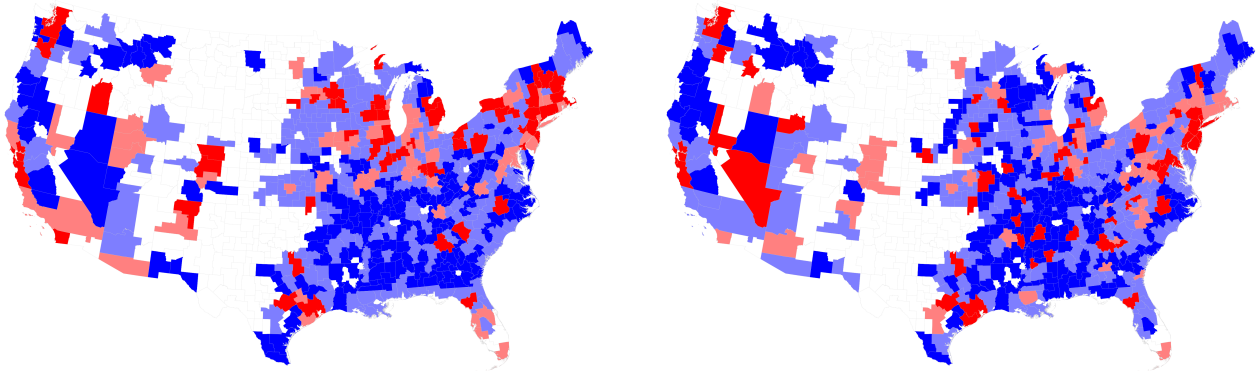
Notes: Scatter plots of log-income per capita in the model (horizontal axis) and data (vertical axis) in 1950 (top-left panel), 1970 (top-right panel), 1990 (bottom-left panel), and 2010 (bottom-right panel). The size of the circles is proportional to population. All variables are displayed as deviations from the mean.

Figure A.10: **Income per capita by location-sector: data and model (untargeted)**



Notes: Bin-scatter plots of log-income per capita in the model (horizontal axis) and data (vertical axis) by location-sector, weighted by individuals in each location-sector, residualized with respect to location and sector fixed effect, in 1990 (left panel) and 2010 (right panel).

Figure A.11: **Future scenarios: Declining frictions to diffusion**

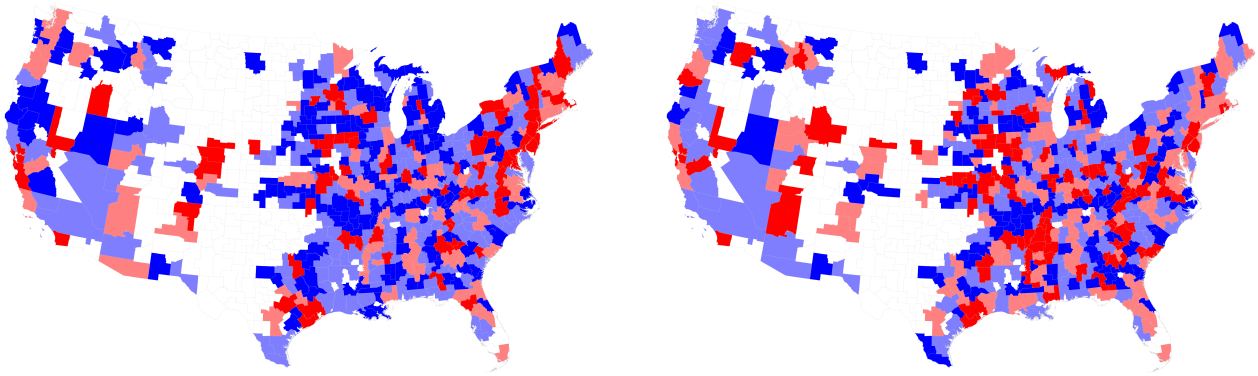


a. Declining geographical frictions

b. Declining technological frictions

Notes: Panel (a) shows log-population in 2050 after a 50% decline in adjusted diffusion frictions ($(d_{(m,r)\to(n,s),t})^{\theta}$) across locations (i.e., for all $m \neq n$), in deviation from a status quo in which frictions are kept at their 2010 values. Panel (b) shows log-population in 2050 after a 50% decline in adjusted diffusion frictions across sectors (i.e., for all $r \neq s$), in deviation from a status quo in which frictions are kept at their 2010 values. Blue CZs correspond to a net population gain (light blue: below-median increase; dark blue: above-median increase), while red CZs correspond to a net population loss (light red: below-median loss; dark red: above-median loss).

Figure A.12: **Future scenarios: Autonomous vehicles**

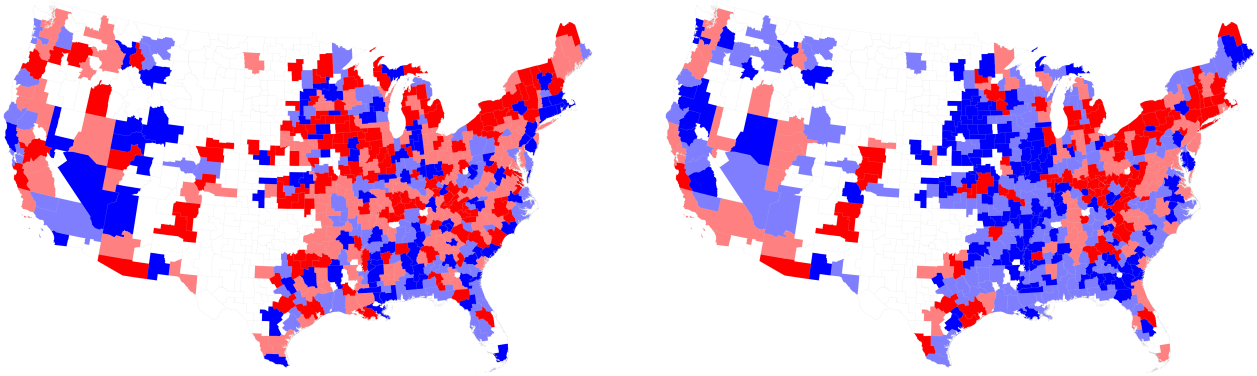


a. Tech. wave shock to Transportation

b. Declining frictions to and from Transportation

Notes: The maps show log-population in 2050 after technological wave shocks of magnitude +20% to B2 (panel a), as well as a 50% decline in adjusted diffusion frictions ($(d_{(m,r)\to(n,s),t})^{\theta}$) from (to) B2 to (from) both G1 and H1 (panel b), in deviation from a status quo in which $\alpha_{s,t}$ are kept at their 2010 values. Blue CZs correspond to a net population gain (light blue: below-median increase; dark blue: above-median increase), while red CZs correspond to a net population loss (light red: below-median loss; dark red: above-median loss).

Figure A.13: **Future scenarios: Pharmaceuticals and Agriculture**



a. Tech. wave shock to Pharmaceuticals

b. Tech. wave shock to Agriculture

Notes: The maps show log-population in 2050 after technological wave shocks of magnitude +20% to A3 (panel a) and to A1 (panel b) in deviation from a status quo in which $\alpha_{s,t}$ are kept at their 2010 values. Blue CZs correspond to a net population gain (light blue: below-median increase; dark blue: above-median increase), while red CZs correspond to a net population loss (light red: below-median loss; dark red: above-median loss).

B Data description

In this section, we provide details on the construction of the data on population, human capital, and employment by industry at the commuting zone level. Our starting points are the publicly available data from the Integrated Public Use Microdata Series (IPUMS, Ruggles et al., 2021) and the National Historical Geographic Information System (NHGIS, Manson et al., 2021), and the restricted full-count censuses for the decades until 1940.[1]

Population at the CZ level until 1940 is constructed using the full-count censuses. In particular, we use the mapping of each Census place to their corresponding coordinates performed by Berkes et al. (2023) to assign each individual in the historical Census to a 1990 commuting zone. Total population is then computed as the total number of individuals whose coordinates fall in that 1990 commuting zone. For the decades 1950 onwards, we use county-level population counts from the NHGIS and build crosswalks from counties to 1990 commuting zones based on the shares of overlapping areas. Using the full-count censuses, we also build crosswalks from historical counties to 1990 commuting zones for the decades until 1940 (which are then used to create measures of human capital and industry composition). For the purposes of the model calibration, since the model cannot rationalize extreme declines in population within one model period (due to the overlapping generations structure), we set the maximum decline in population between any two periods to -30%. The resulting total correction is negligible, with a maximum of 0.038% of the total population in 1930.

Measures of the local density of human capital are assembled starting from county-level data from the NHGIS and aggregated at the level of 1990 commuting zones using the same crosswalks. The specific variables used to construct the measure vary by decade depending on the availability. We make the measures comparable across decades by converting them into the corresponding ranking. In 1890, we interpolate measures from the 1880 and 1990 decennial census. The 1880 measure concerns the share of people who attended school. Between 1900 and 1930, the measure represents the (inverse of the) share of illiterate people. In 1950, the measure reflects the median years of schooling of the population. From 1970 onwards, the measure corresponds to the share of population with at least a college degree.

Local industry composition is constructed using the full-count census (until 1940) and the IPUMS (from 1950 onwards). We consider the number of people in each of the 12 main industries in the 1950 Census Bureau industrial classification system. To allocate individuals to 1990 commuting zones, we construct area-based crosswalks from State Economic Areas (in 1950), County Groups (in 1970), and PUMAs (in 1990 and 2010).

[1]The data from the 1890 decennial census are not available; hence, we construct the 1890 observations by linearly interpolating the observations from the 1880 and 1900 decennial censuses. For the 2010 observation, we use multi-year averages of the American Community Survey (ACS).

We construct data on average income at the commuting zone level for the decades 1950, 1970, 1990, and 2010. For the 1950 decade, we use binned median income at the level of the 1950 county. We use a mid-point assumption to assign average values to each bin, and use a series of crosswalks to construct 1950 income for each 1990 commuting zone (this measure can be interpreted as a weighted average of the median income of the 1950 counties that constitute each 1990 commuting zone). For the 1970 decade, we compute CZ-level income using data on the number of individuals by 1970 county within each bin of income (e.g., \$1-\$999, \$1,000-\$1,999, etc.). Again, we use the mid-point method to assign average values for each bin, and a series of crosswalks to map 1970 counties to 1990 commuting zones. CZ-level income for 1990 and 2010 is computed as average per capita income at the county level, weighted by the share of each county in the corresponding commuting zone.

Finally, we construct data on average income at the sector-CZ level for 1990 and 2010. For 1990 and 2010, IPUMS provides information on individual income and industry of employment, following the Census industry classification. We first use a Census-provided crosswalk to map each Census industry code into one or more 3-digit NAICS codes. We then use the crosswalk between NAICS codes and IPC classes built by Lybbert and Zolas (2014) to map each individual in the IPUMS to a distribution of technology class-groups. Since IPUMS locates individuals at the PUMA level, we also map them probabilistically to 1990 commuting zones using the crosswalk built by Berkes and Gaetani (2023).

C Equilibrium definition

Formally, an equilibrium in the simplified model of Section 3 is defined as follows:

Definition C.1. *Given*

$$L_0, \{\lambda_{n,s,0}\}_{(n,s)\in N\times S},\ \{u_n\}_{n\in N}, \{d_{(m,r)\to(n,s)}\}_{(m,r),(n,s)\in(N\times S)^2},$$

and a path of exogenous variables

$$\{f_t\}_{t\geq 0},\ \{\alpha_{r,t}\}_{r\in S, t\geq 0},\ \{\epsilon_{n,s,t}\}_{n,s\in N\times S,\, t\geq 0},$$

an equilibrium is a path for the endogenous variables

$$\{\lambda_{n,s,t},\ \pi_{n,s,t},\ L_{n,s,t}\}_{n,s\in N\times S,\, t\geq 0}$$

that satisfies the following conditions:

1. *Migration and occupational probabilities* $\{\pi_{n,s,t}\}_{n,s\in N\times S,\, t\geq 0}$ *satisfy Equation* (13).
2. *The path for* $\{\lambda_{n,s,t}\}_{n,s\in N\times S,\, t\geq 0}$ *satisfies the law of motion of Equation* (9).
3. *Population by location-sector* $\{L_{n,s,t}\}_{n,s\in N\times S,\, t\geq 0}$ *satisfies the transition identity* (14).

D Details on model extensions for quantitative analysis

In this section, we provide more details on the model extensions for quantitative analysis introduced in Section 4.1, with a specific focus on the equilibrium conditions and the log-linearized model dynamics. To make this section self-containing, in what follows we summarize the extensions and reproduce the main equations.

Individuals live for three periods ("child", "young adults", and "old adults"). Every child is born in the location of their parents. At the end of childhood, the agent makes their migration and occupational choice by selecting which location n they migrate to and which sector s they specialize in. This choice is irreversible, so each agent spends the youth and old periods in the same location-sector. The productivity of agent i who is young at time t and old at time $t+1$ is denoted by $q^y_{n,s,i,t}$ and $q^o_{n,s,i,t+1}$, respectively. Agents are endowed with one unit of inelastically-supplied labor per period.

Denoting by $L^y_{n,s,t}$ and $L^o_{n,s,t}$ the mass of young and old agents, the following identity holds:

$$L^o_{n,s,t} \equiv L^y_{n,s,t-1}.$$

After the migration and occupational choices have been made, each young adult in period t has f_t children. Denoting by $L^k_{n,t}$ the mass of children born in location n at time t, we have that

$$L^k_{n,t} \equiv f_t \sum_{s \in S} L^y_{n,s,t}.$$

Migration and occupational decisions maximize expected lifetime utility, subject to migration costs and idiosyncratic utility draws. Individual i born in location m and selecting location-sector (n, s) has utility given by:

$$U_{m \to n}(u_{n,t}, x_{n,s,i}, c^y_{n,s,i,t}, c^o_{n,s,i,t+1}) = u_{n,t} \frac{x_{n,s,i} \left(c^y_{n,s,i,t}\right)^{\beta} \left(c^o_{n,s,i,t+1}\right)^{1-\beta}}{\mu_{m \to n}}, \tag{D.1}$$

where $u_{n,t}$ is the level of amenities in city n at time t, $x_{n,s,i}$ is an idiosyncratic utility draw from a Fréchet distribution with shape parameter $\zeta > 1$, $\mu_{m \to n}$ represents migration costs of moving from m to n, $c^y_{n,s,i,t}$ and $c^o_{n,s,i,t+1}$ denote consumption in the youth and the old period, and $\beta \in (0, 1)$ is the weight of youth consumption in lifetime utility. There are no markets to smooth consumption over time and generations. Therefore, in every period, individual consumption is equal to production.

Total output in the economy is given by a linear aggregator of individual productivities across

all locations and sectors:

$$Y_t = \sum_{n \in N} \sum_{s \in S} \left(L^y_{n,s,t} \mathbb{E}[q^y_{n,s,\cdot,t}] + L^o_{n,s,t} \mathbb{E}[q^o_{n,s,\cdot,t}] \right),$$

where $\mathbb{E}[q^y_{n,s,\cdot,t}]$ ($\mathbb{E}[q^o_{n,s,\cdot,t}]$) denotes the average productivity of young (old) agents in location-sector (n, s), and $L^y_{n,s,t}$ ($L^o_{n,s,t}$) denotes the corresponding mass of individuals.

Young adults undergo the imitation or innovation process outlined in Section 3.1.3 to determine their youth productivity $q^y_{n,s,i,t}$ upon moving. Note that the relevant variables ($\epsilon_{n,s,t}$ and $\alpha_{n,s,t}$) are known at the time of the migration and occupational choice. Under Assumption A1, the local distribution of productivity among young agents remains Fréchet and the corresponding scale parameter, $\lambda_{n,s,t}$, follows the law of motion in Equation (9). Old adults retain the same youth-period idiosyncratic productivity rescaled by a factor A^o which captures the experience productivity premium enjoyed by old relative to young agents:

$$q^o_{n,s,i,t+1} = A^o q^y_{n,s,i,t}.$$

As common in the quantitative economic geography literature, local amenities are assumed to be an isoelastic function of local population:

$$u_{n,t} = \bar{u}_{n,t} \, L^{\omega}_{n,t},$$

where $\bar{u}_{n,t}$ is the exogenous time-varying component of local amenities, and ω is the elasticity of local amenities to population, that can account for both congestion ($\omega < 0$) and agglomeration ($\omega > 0$) forces.

The probability of an individual born in location m to select city-sector (n, s) is

$$\pi_{m \to (n,s),t} = \frac{\left(u_{n,t} \frac{\lambda_{n,s,t}}{\mu_{m \to n}} \right)^{\zeta}}{\sum_{l,r} \left(u_{l,t} \frac{\lambda_{l,r,t}}{\mu_{m \to l}} \right)^{\zeta}}. \tag{D.2}$$

Notice that since the experience premium A^o scales productivity in all locations by the same factor, it does not directly affect migration probabilities.

The following identity between children and young adults holds for all cities and sectors:

$$L^y_{n,s,t} \equiv \sum_{m=1}^{N} \pi_{m \to (n,s),t} L^k_{m,t-1}.$$

Consider now the dynamics of the migration shares, $\pi_{n,t}$, in response to an arbitrary deviation

of $\lambda_{n,s,t}$ from the BGP, denoted by $\hat{\lambda}_{n,s,t}$ (for simplicity, assume here that local amenities are exogenous, i.e., $\omega = 0$). Denoting by $\pi_{m\to n,t}$ the probability of migrating from m to n, and log-linearizing this probability around the BGP, yields

$$\hat{\pi}_{m\to n,t} = \zeta \sum_{s\in S} \left\{ (1-\pi^*_{m\to n})\pi^*_{s|n}\hat{\lambda}_{n,s,t} - \sum_{l\neq n} \pi^*_{m\to(l,s)}\hat{\lambda}_{l,s,t} \right\}, \tag{D.3}$$

where $\pi^*_{s|n}$ is the BGP probability of choosing sector s conditional on migrating to n (note that this probability does not depend on the city of origin m). The total migration probability to location n can be written as

$$\pi_{n,t} = \sum_{m\in N} \pi_{m\to n,t}\pi_{m,t-1}.$$

Assuming the economy is in BGP at $t-1$, the response of the local migration share, $\pi_{n,t}$, can be written as

$$\hat{\pi}_{n,t} = \sum_{m\in N} \frac{\pi^*_{m\to n}\pi^*_m}{\pi^*_n}\hat{\pi}_{m\to n,t}, \tag{D.4}$$

where $\hat{\pi}_{m\to n,t}$ is given by Equation (D.3). The term $\frac{\pi^*_{m\to n}\pi^*_m}{\pi^*_n}$ corresponds to the probability that a youth living in n was born in m. This probability can be interpreted as the "reliance" of n on migration from m. Intuitively, in the presence of migration frictions, local population growth in n is determined by local productivity growth relative to productivity growth in other locations, where cities with higher migration flows to and from n have more weight than cities with lower migration flows. Equation (D.3), combined with Equation (D.4), is equivalent to the expression for changes in regional labor supply obtained by Borusyak et al. (2022), who study the role of migration frictions in shaping the spatial response to local shocks.

E Details on the calibration procedure

E.1 Gravity equation for migration flows

Consider first the gravity equation for migration probabilities (26):

$$\log(\pi_{m\to n,t}) = \psi^0_{m,t} + \psi^1_{n,t} - \zeta\bar{\mu}^0\mathbf{1}_{m\neq n} - \zeta\bar{\mu}^D D_{m,n}.$$

To derive this equation, take the sum across all sectors s of the right-hand-side of Equation (D.2) to obtain:

$$\pi_{m\to n,t} = \frac{\sum_s \left(u_{n,t}\frac{\lambda_{n,s,t}}{\mu_{m\to n}}\right)^\zeta}{\sum_{l,r}\left(u_{l,t}\frac{\lambda_{l,r,t}}{\mu_{m\to l}}\right)^\zeta}.$$

Taking logs on both sides and using the definition of the migration costs in Equation (25) yields the gravity representation in Equation (26).

We estimate this relationship using data from the 1990 IPUMS to recover the composite parameters $\zeta\bar{\mu}^0$ and $\zeta\bar{\mu}^D$. We restrict the sample to "young adult" individuals in the IPUMS (i.e., between the age of 20 and 40), and assign each individual to a city of residence through a probabilistic crosswalk (based on areas) from Consistent Public-Use Metro Area (CONSPUMA) to 1990 commuting zones. Similarly, we assign each individual to a city of birth through a probabilistic crosswalk from state of birth to 1990 commuting zones based on 1970 population.

Table E.5: **Gravity equation for migration flows**

	Dependent var.: Migration probability
Distance in 1,000 km	-2.8124***
	(0.1857)
Origin CZ ≠ Destination CZ	-0.7888***
	(0.1204)
Origin location FE	yes
Destination location FE	yes
# Obs.	235,225
Estimation	PPML

Notes: PPML estimates. The sample includes all pairs of 485 commuting zones in our main sample. The dependent variable is the migration probability between each pair of commuting zones. Standard errors clustered at the commuting zone of origin and the commuting zone of destination in parenthesis. ***$p < 0.01$.

We estimate Equation (26) via Poisson Pseudo Maximum Likelihood (PPML), which is useful

to accommodate zero values (there are 141 commuting zone pairs with migration probability equal to zero). Results are reported in Table E.5. The estimate of the composite parameter, $\zeta\bar{\mu}^D$, is 2.81, implying that increasing distance by 100 km decreases the migration probability by roughly 24% which, at the average migration probability, corresponds to 0.05 percentage points.

E.2 Recovering the path of local productivities ($\lambda_{n,s,t}$)

To calibrate the path of the scale parameters $\lambda_{n,s,t}$, we postulate (and later validate) a simple notion of patenting in the model that allows us to draw a transparent connection between patent data and the path of local productivities. In particular, we assume that young individuals, when undergoing their imitation or innovation process outlined in Section 3.1.3, will file for a patent for their best innovation draw if it results in an idea of quality above a time-varying threshold, Q_t.

Notice that the best innovation draw (the second term in Equation 7) is distributed Fréchet with shape parameter θ and scale parameter, denoted by $\Lambda_{n,s,t}$, which is defined as

$$\Lambda_{n,s,t} \equiv \sum_{m\in N}\sum_{r\in S}\lambda_{m,r,t-1}\left(\frac{\epsilon_{n,s,t}\alpha_{r,t}}{d_{(m,r)\to(n,s),t}}\right)^{\theta} = \lambda_{n,s,t} - \lambda_{n,s,t-1},$$

where the last equality follows from Equation (9).

The probability that a young individual from (n,s) patents at time t is equal to the probability that their best innovation draw is above the threshold Q_t, or equivalently:

$$\frac{Pat_{n,s,t}}{L^y_{n,s,t}} = 1 - e^{-\Lambda_{n,s,t}Q_t^{-\theta}}, \tag{E.5}$$

where $Pat_{n,s,t}$ denotes the total number of patents filed at time t in location-sector (n,s).[2] Since $Pat_{n,s,t}/L^y_{n,s,t}$ is a typically a small number (in the order of 10^{-3}), the following expression provides an accurate approximation for Equation (E.5):[3]

$$\frac{Pat_{n,s,t}}{L^y_{n,s,t}} \approx \Lambda_{n,s,t}Q_t^{-\theta}. \tag{E.6}$$

Given values for θ and Q_t, we can then set the path of $\lambda_{n,s,t}$ and $u_{n,t}$ so to satisfy Equation (E.6) for all (n,s,t) and, simultaneously, match total population by location in all time periods.[4]

[2]Notice that Equation (E.5) rationalizes instances where $Pat_{n,s,t}$ is equal to zero as $\epsilon_{n,s,t} = 0$.

[3]To see this, rearrange Equation (E.5) and take logarithms on both sides to obtain $\log\left(1 - \frac{Pat_{n,s,t}}{L^y_{n,s,t}}\right) = -\Lambda_{n,s,t}Q_t^{-\theta}$. Since $Pat_{n,s,t}/L^y_{n,s,t}$ is typically in the order of 10^{-3}, Equation (27) provides a highly accurate approximation for this expression.

[4]This inversion is feasible as long as $Pat_{n,s,t} > 0$ for at least one s in all locations in the initial BGP. To minimize instances in which the inversion is not feasible, we compute $Pat_{n,s,t}$ for the initial BGP (1890) as the sum of all patents in location-sector (n,s) filed since the beginning of the sample and the 1890 decade. Since these instances

E.3 Gravity equation for knowledge flows

Consider now the gravity equation for knowledge flows (Equation 29) introduced and derived in Section 4.2.3:

$$\log(\eta_{(m,r)\to(n,s),t}) = \phi^0_{m,r,t} + \phi^1_{n,s,t} - \theta\delta^G \mathbf{1}_{m\neq n} - \theta\delta^D D_{m,n} - \theta\delta^M M_{m\to n,t} - \theta\delta^K_{r\to s}.$$

To estimate the composite parameters $\theta\delta^G$, $\theta\delta^D$, $\theta\delta^M$, and $\theta\delta^K_{r\to s}$, we leverage our patent citation data. We include only citations added by applicants. We restrict the sample to patents filed since 1980 and issued since 2000 (patents began reporting inventor-added citations separately from examiner-added ones only in 2000, see Alcacer and Gittelman, 2006). We compute $\eta_{(m,r)\to(n,s),t}$ as the share of citations given by patents in (n,s) and directed to patents in (m,r).

Every citing patent in our regression has a total weight of one. In other words, every citation is weighted by the inverse of the total number of citations given by the citing patent. To account for the fact that knowledge flows are more likely to be tacit and less likely to be captured by citations, we add an artificial citation to each patent's list of references to a local patent whose technology classes are identical to the citing patent. This assumption amounts to assuming that the weight of local tacit flows is inversely proportional to the number of inventor-provided citations. This also guarantees that all patents, even those without backward citations, are included in the estimation.

We estimate this relationship by PPML and report results in Table E.6. In addition to the composite parameters $\theta\delta^G$, $\theta\delta^D$, and $\theta\delta^M$, we also obtain a full set of bilateral transmission costs across sectors ($\delta^K_{r\to s}$), which we show in the heatmap of Figure E.14.

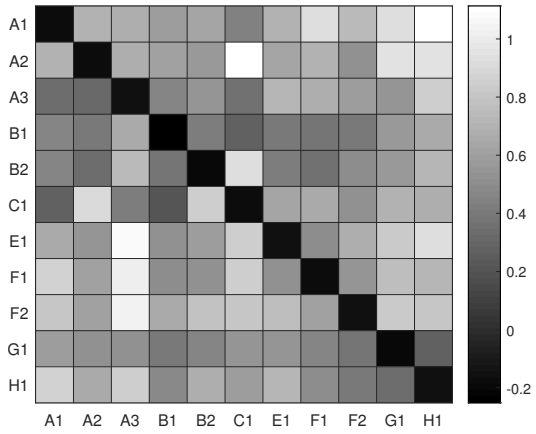
are rare and only involve small CZs, our results are not sensitive to alternative ways of dealing with zeroes. We choose units of the final good so that the geometric average of $\lambda_{n,s}$ is equal to one in the first time period.

Table E.6: **Gravity equation for knowledge flows**

	Dependent var.: Share of citations
Origin CZ ≠ Destination CZ	-6.4261***
	(0.0189)
Distance in 1,000 km	-0.1645***
	(0.0087)
Migration exposure	3.3359***
	(0.0772)
Origin location-sector FE	yes
Destination location-sector FE	yes
Origin-Destination sector FE	yes
# Obs.	27,527,528
Estimation	PPML

Notes: PPML estimates. The sample includes patents filed since 1980 and issued since 2000. Observations are all the combinations of pairs of location-sectors. The dependent variable is the share of citations given by each destination location-sector to each origin location-sector, where each citing patent is given a weight of one. Class-groups are described in Appendix Table A.1. Standard errors clustered at the destination location-sector in parenthesis. ***$p < 0.01$.

Figure E.14: **Knowledge transmission costs across sectors**



Notes: PPML estimates of $\delta^K_{r\to s}$, from regression of Appendix Table E.6. The sample includes patents filed since 1980 and issued since 2000. Observations are all the combinations of pairs of location-sectors. The dependent variable is the share of citations given by each destination location-sector to each origin location-sector, where each citing patent is given a weight of one. Rows correspond to citing (idea destination) sectors. Columns correspond to cited (idea origin) sectors. Number of observations: 27,527,528. Class-groups are described in Appendix Table A.1.

F Robustness

In this section, we explore robustness of the model's main results to two choices of parameters explained in Section 4. In particular, we consider values of ζ (the elasticity of migration with respect to average income, which is set to 4 in the baseline) and ω (the elasticity of residential amenities to population, which is set to -0.15 in the baseline) in the range considered by Allen and Donaldson (2022). In particular, we set ζ and ω to the minimum ($\zeta = 2$ and $\omega = -0.392$, respectively) and maximum ($\zeta = 6$ and $\omega = -0.039$, respectively) values considered by Allen and Donaldson (2022).

Panel A and B of Table F.7 report the results of the main counterfactuals (as in Table 3), for these alternative assignment of ζ and ω, respectively. The results are not substantially affected by these parameter choices. As expected, alternative values of these elasticities produce some differences in the quantitative effects, with the share of the correlation explained by our mechanism ranging from 88.1% when congestion externalities are set at the minimum in the range of values estimated by Allen and Donaldson (2022) ($\omega = -0.039$) to 41.7% when congestion externalities are set at their maximum ($\omega = -0.392$). The decomposition of the effect between geographical and technological frictions (comparing columns 2-4 to columns 5-7) remains qualitatively consistent across parametrizations, although with some quantitative variation, ranging from 44.1% when $\zeta = 6$ to 69.0% when $\zeta = 2$.

Table F.7: **Population growth and technological wave shocks: robustness**

	Dependent var.: Growth rate of population under						
	Full model	Model without tech. waves					
					No diffusion across fields		
Panel A: Robustness with respect to ζ							
		Main	$\zeta = 2$	$\zeta = 6$	Main	$\zeta = 2$	$\zeta = 6$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposure to tech. waves ($Exp_{n,t}$)	0.341***	0.119	0.169*	0.104	0.237**	0.288***	0.209**
	(0.099)	(0.100)	(0.100)	(0.099)	(0.098)	(0.099)	(0.098)
Difference from empirical coefficient	-	0.222	0.172	0.237	0.104	0.053	0.133
Share explained by tech waves ($\alpha_{s,t}$)		65.1%	50.4%	69.5%	30.5%	15.6%	38.8%
Decomposition:							
- Share explained by tech. frictions					46.9%	31.0%	55.9%
- Share explained by geo. frictions					53.1%	69.0%	44.1%
Calibrated θ	4.48	4.48	4.45	4.50	4.48	4.45	4.50
# Obs.	2,852	2,852	2,852	2,852	2,852	2,852	2,852
R^2	0.54	0.54	0.54	0.54	0.54	0.54	0.54
Panel B: Robustness with respect to ω							
		Main	$\omega = -0.392$	$\omega = -0.039$	Main	$\omega = -0.392$	$\omega = -0.039$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposure to tech. waves ($Exp_{n,t}$)	0.341***	0.119	0.199*	0.041	0.237**	0.278***	0.195*
	(0.099)	(0.100)	(0.099)	(0.101)	(0.098)	(0.098)	(0.099)
Difference from empirical coefficient	-	0.222	0.142	0.301	0.104	0.063	0.147
Share explained by tech waves ($\alpha_{s,t}$)		65.1%	41.7%	88.1%	30.5%	18.5%	43.0%
Decomposition:							
- Share explained by tech. frictions					46.9%	44.3%	48.8%
- Share explained by geo. frictions					53.1%	55.7%	51.2%
Calibrated θ	4.48	4.48	4.48	4.48	4.48	4.48	4.48
# Obs.	2,852	2,852	2,852	2,852	2,852	2,852	2,852
R^2	0.54	0.54	0.54	0.54	0.54	0.54	0.54
Log-population (lags 1 and 2)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	CD×T	CD×T	CD×T	CD×T	CD×T	CD×T	CD×T

Notes: CZ-level regression, 1910-2010. Dependent variable defined as growth rate of population over 20 years. "CD×T" denotes Census Division-time fixed effects. Standard errors clustered at the CZ and the Census Division-time level in parenthesis. Exposure to the technological wave is defined as in Equation (1). Columns 2-4 display the counterfactual in which $\alpha_{s,t}$ are kept constant at their 1890 BGP values. Column 5-7 display the counterfactual in which knowledge flows are restricted to within-field flows only. Controls include log-total patents, human capital, and industry composition. *$p < 0.01$, **$p < 0.05$, ***$p < 0.01$.

G Deriving the specialization measure

To rationalize the measure of specialization in Equation (31), consider the simple model of Section 3, in which we impose the following assumption on the distribution of technological wave shocks:

Assumption A5. *Technological wave shocks are uncorrelated across sectors and have a constant variance:*

1. $Cov(\hat{\alpha}_{s,t}, \hat{\alpha}_{r,t}) = 0$ *for all* $s \neq r$
2. $Var(\hat{\alpha}_{s,t}) = V$ *for all* $s \in S$.

Using Assumption A5 in combination with Assumption A2.2, we derive the following theoretical result, that links the volatility of local population growth to the local degree of specialization:

Proposition G.1. *Under Assumptions A2.2 and A5, the variance of the percentage change in the population share of location n satisfies:*

$$Var(\hat{\pi}_{n,t}) \propto \sum_{r \in S} \left(\eta^*_{r \to n} - \bar{\eta}^*_r\right)^2. \tag{G.7}$$

Proof. Factoring out $(1 - \pi^*_n)$ from Equation (20), and realizing that $\pi^*_{s|-n} \equiv \sum_{m \neq n} \frac{\pi^*_{m,s}}{1-\pi^*_n}$, we can rewrite:

$$\hat{\pi}_{n,t} \propto (1 - \pi^*_n) \sum_{r \in S} \left\{ \sum_{s \in S} \pi^*_{s|n} \eta^*_{r \to (n,s)} - \sum_{s \in S} \sum_{m \neq n} \frac{\pi^*_{n,s}}{1 - \pi^*_n} \eta^*_{r \to (m,s)} \right\} \hat{\alpha}_{s,t}. \tag{G.8}$$

Under Assumption A5, the technological wave shocks $\alpha_{s,t}$ have zero covariance and common variance V. In this case, the variance of $\hat{\pi}_{n,t}$ is equal to

$$Var(\hat{\pi}_{n,t}) \propto (1 - \pi^*_n)^2 \sum_{r \in S} \left(\eta^*_{r \to (n,s)} - \bar{\eta}^*_{r \to (-n,s)}\right)^2,$$

where $\eta^*_{r \to n} \equiv \sum_{s \in S} \pi^*_{s|n} \eta^*_{r \to (n,s)}$, and $\bar{\eta}^*_{r \to (-n,s)} \equiv \sum_{s \in S} \sum_{m \neq n} \frac{\pi^*_{n,s}}{1-\pi^*_n} \eta^*_{r \to (m,s)}$.

Under Assumption A2.2, all cities are negligible in size compared to the overall economy. Hence, under Assumptions A2.2 and A5, the variance of the percentage change in the population share of location n is equal to the measure of specialization in Equation (31):

$$Var(\hat{\pi}_{n,t}) \propto \sum_{r \in S} \left(\eta^*_{r \to n} - \bar{\eta}^*_r\right)^2 \equiv Spec_n.$$

□

References

Alcacer, J. and M. Gittelman (2006): "Patent citations as a measure of knowledge flows: The influence of examiner citations," *The review of economics and statistics*, 88, 774–779.

Allen, T. and D. Donaldson (2022): "Persistence and path dependence in the spatial economy," Tech. rep.

Berkes, E. and R. Gaetani (2023): "Income segregation and the rise of the knowledge economy," *American Economic Journal: Applied Economics*, 15, 69–102.

Berkes, E., E. Karger, and P. Nencka (2023): "The census place project: A method for geolocating unstructured place names," *Explorations in Economic History*, 87, 101477.

Borusyak, K., R. Dix-Carneiro, and B. Kovak (2022): "Understanding Migration Responses to Local Shocks," *Available at SSRN 4086847*.

Lybbert, T. J. and N. J. Zolas (2014): "Getting patents and economic data to speak to each other: An 'algorithmic links with probabilities' approach for joint analyses of patenting and economic activity," *Research Policy*, 43, 530–542.

Manson, S., J. Schroeder, D. Van Riper, T. Kugler, and S. Ruggles (2021): "IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS," http://doi.org/10.18128/D050.V16.0.

Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek (2021): "IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021," https://doi.org/10.18128/D010.V11.0.