Technological Waves and Local Growth*

Enrico Berkes† Ruben Gaetani‡ Martí Mestieri§

8th September 2023

Abstract

We develop a spatial model of endogenous growth via frictional knowledge diffusion to examine the effect of technological waves—defined as long-term shifts in the importance of specific knowledge fields—on local population 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 more than half 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 technological scenarios may have large geographical effects.

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

*A previous version of this paper circulated as "Cities and Technological Waves." We thank Mike Andrews, Paco Buera, Klaus Desmet, Ed Glaeser, Xian Jiang, Ben Jones, Sara Mitchell, Joel Mokyr, Nicola Persico, Frédéric Robert-Nicoud, Bruce Weinberg, and seminar attendees at University of Colorado Boulder, University of Toronto, Ohio State University, University of Pittsburgh, Georgetown University, University of Geneva, Conference of Swiss Economists Abroad, 2020 NBER Summer Institute (Urban Economics), VMACS Junior Conference, 2021 AEA Meetings, 2021 European Meeting of the Urban Economics Association, 2021 Canadian Summer Conference in Real Estate and Urban Economics, 2021 Barcelona GSE Summer Forum (Firms in the Global Economy), 2021 SED Meeting, 2021 North American Meeting of the Urban Economics Association, 2022 European Summer Meeting of the Econometric Society, and 2022 Macro Day at the University of Southern California for helpful comments and discussions. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System. All errors are our own.

†University of Maryland, Baltimore County, enrico.berkes@umbc.edu.
‡University of Toronto, ruben.gaetani@utoronto.ca.
§UPF, CREi, BSE, Federal Reserve Bank of Chicago, 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). Empirical research demonstrates that idea diffusion is localized so that geographical proximity is a key determinant of people's capacity to learn and adopt existing knowledge (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, as exemplified in the United States 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). However, the economic channels underlying these patterns remain largely unexplored. On the one hand, progress on this issue has been hampered by the lack of comprehensive data on the geography of innovation that span a sufficiently long period. On the other hand, identifying specific channels has proven difficult both empirically and theoretically, due to difficulties 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 develops a new theory that allows us to investigate the link between the evolution of the technological environment and the rich dynamics of local growth. Using a new 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"—experience systematically higher growth over the long-term.

Motivated by this finding, we develop a model that combines a spatial equilibrium setting with a theory of economic growth that emphasizes the role of knowledge recombination, imitation, and diffusion. The model formalizes the relationship between technological waves and the growth dynamics of cities and illustrates how frictions in the diffusion of ideas mediate such relationship. The model remains highly tractable, from a theoretical and quantitative viewpoint, allowing us to isolate the channels at play and estimate the contribution of the underlying economic forces. The quantitative results suggest that frictions to idea diffusion explain more than half of the reduced-form relationship between exposure to technological waves and local population growth. The quantified model can endogenously account for prominent features of the U.S. economic geography, such as the decline of cities in the Central States throughout the last century and the success of many of the modern innovation hubs.

To begin our analysis, we use our historical patent data to show that, over the last century, cities' growth trajectories were systematically affected by changes in the technological environment.

Our data reveals that the distribution of innovation activities across fields varies significantly between cities. In addition, the prominence of various fields of knowledge 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 innovation activities initially concentrated in expanding fields experience systematically higher population growth than cities with innovation activities concentrated in declining fields. We also document that knowledge diffusion, measured by patent citations, is persistently localized, in both geographical and technological spaces. 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 rise of some cities and the decline of others. In the remainder of the paper, we formalize this mechanism in a spatial model of endogenous growth with innovation and frictional idea diffusion and then use it to quantify our proposed mechanism.

In the model, agents make migration and occupational decisions based on their expected lifetime productivity in each location and sector. Productivity is determined by a decision to imitate or innovate. Agents can either imitate an idea drawn from the local knowledge distribution or innovate by enhancing an idea drawn from the distribution of any location and sector. The applicability of an idea is affected by frictions that reflect both geographical and technological distance. These frictions imply that knowledge drawn *within* any location-sector can be converted into new inventions more efficiently than knowledge drawn from other locations and sectors. Because of this, a city's knowledge stock affects its current productivity and future innovation potential, making the local growth trajectory sensitive to technological wave shocks. These shocks affect the returns on innovation in different fields in a manner that favors locations with a greater availability of complementary ideas.

The framework remains tractable for any arbitrary 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 Frećhet structure, which implies an intuitive law of motion for its scale parameter and hence its average productivity. Additionally, the Fréchet structure enables us to characterize knowledge flows in closed form via a gravity representation that can be estimated using patent citation data.

To investigate 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 specific 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.

We then return to the full non-linear framework to evaluate quantitatively the proposed mechanism. We extend the model to include empirically relevant dimensions such as overlapping generations, migration costs, and endogenous congestion forces. The model has a recursive structure that allows us to calibrate the parameters and recover the unobserved disturbances, including technological wave shocks, by imposing a small set of transparent assumptions. The calibrated model successfully captures key moments of interest that are not directly targeted, such as the relationship between city size and average city income.

Our quantitative 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 more than half of the reduced-form relationship between technological waves and local growth. A decomposition reveals that frictions to knowledge diffusion across geographical areas and technological fields are both significant, with the former accounting for approximately one-third of the overall effect and the latter for approximately two-thirds. The quantification indicates that technological waves throughout the last century had heterogeneous effects on local growth, contributing to delineate the current economic geography of the United States. For example, technological waves, via the endogenous process of frictional idea diffusion, penalized local growth in cities in the Central States, and positively contributed to the growth of modern innovation hubs.

The endogenous mechanism of frictional diffusion 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. This effect reflects two separate forces. First, frictions in the diffusion of knowledge across fields imply that in response to technological wave shocks, productivity growth is greater in some sectors than others. Negative shocks to some sectors are likely to be compensated by positive shocks to other sectors, thereby reducing the chance of large swings (either positive or negative) in the productivity growth of cities with greater diversification. Second, frictions to diffusion across geographical areas imply that more diversified cities have a broader pool of ideas from which to draw, thereby mitigating the impact of (positive or negative) shocks to individual sectors on the evolution of local productivity.

Finally, we use the quantitative model to explore the predicted local growth dynamics in the coming decades under various technological scenarios. In particular, we examine which cities benefit in terms of population growth—and which do not—relative to the status quo in the following

scenarios: (1) an increase in the importance of transportation-related technologies due to the emergence of new modes of transportation such as autonomous vehicles; (2) an increase in the centrality of pharmaceuticals and biotech in response to new challenges in global health; (3) a comeback of agriculture as a pivotal sector in the innovation landscape as a result of regulatory changes and increasing demand for sustainable farming. 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 technological and geographical frictions in knowledge diffusion playing a 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), the role of localized knowledge in long-run city dynamics is still largely unexplored.

The complexity of modeling idea diffusion in a spatial setting has been one of the primary obstacles to this evaluation, but two thriving bodies of literature have provided methodological advancements that make this problem more tractable in recent years. First, many papers have developed tractable endogenous growth models that highlight knowledge recombination, imitation, and diffusion as the primary drivers of aggregate productivity growth (e.g., Lucas and Moll, 2014; Perla and Tonetti, 2014; Buera and Oberfield, 2020; Huang and Zenou, 2020). Second, a substantial body of work on quantitative spatial economics has developed tools for analyzing the spatial distribution of economic activity, both within cities (e.g., Ahlfeldt et al., 2015; Heblich et al., 2020) and in a system of locations (e.g., Allen and Arkolakis, 2014; Desmet et al., 2018b).[1] This paper combines insights from these two branches of the literature and develops a parsimonious endogenous growth model in a spatial economy that is highly tractable and 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 the primary determinant of the economic geography (e.g., Davis and Dingel, 2019). This study connects these sets of papers by quantitatively assessing the importance of frictions to idea diffusion for city dynamics. This paper is also related to recent work on the relationship between spatial and aggregate growth, including Cai et al. (2022), who examine spatial growth in a model with idea diffusion through trade and migration, and Arkolakis et al. (2020) and Burchardi et al. (2020),

[1] Buera and Lucas (2018) provide a comprehensive review of the body of literature on models of endogenous growth with idea flows, and Redding and Rossi-Hansberg (2017) provide a comprehensive review of the body of literature on quantitative spatial equilibrium models.

who develop models of spatial growth to quantify the contribution of international migration to aggregate growth.

There is a vast literature investigating the forces governing the long-run evolution of the economic geography, particularly in its tendency to exhibit path dependence and occasional reversals of fortune (e.g., Brezis and Krugman, 1997; Davis and Weinstein, 2002; Bleakley and Lin, 2012; Kline and Moretti, 2014; Allen and Donaldson, 2020), and in its responsiveness to aggregate, regional, or sectoral shocks (e.g., Desmet et al., 2018a; Caliendo et al., 2018). Many recent studies have investigated the specific mechanisms that control the spatial response to local or aggregate shocks, with an emphasis on migration frictions (Hornbeck and Moretti, 2018; Borusyak et al., 2022), commuting across locations (Monte et al., 2018), and general equilibrium effects (Adao et al., 2020). Our paper contributes to this body of research by highlighting an alternative channel underlying the spatial response to shocks, namely the existence of frictions in knowledge transmission. 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 perspective (e.g., Bostic et al., 1997; Simon and Nardinelli, 2002; Michaels et al., 2012; Desmet and Rappaport, 2017) and from a structural and quantitative viewpoint (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) versus urban diversity (Jacobs, 1969) and their effects on city growth. Notable 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.[2] This paper proposes and quantifies a new mechanism by which urban diversification affects long-term city growth by influencing a city's responsiveness to changes in the surrounding technological landscape.[3] In this way, 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 introduces the model and derives the main 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.

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

[3] 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.

2 Data and stylized facts

Our geographical unit of analysis is the 1990 commuting zone (CZ), which we keep fixed over time. Assuming a stable geography allows us to abstract from the pervasive nineteenth and twentieth century phenomena of annexations and redefinition of town borders. Throughout the paper, we refer to cities and CZs interchangeably.

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% per year. From the CUSP, we gather information regarding technology classes, filing date, and location of each inventor listed on a patent. 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. This is the first study to exploit the entire geolocated patent time series provided by the CUSP. Patents with multiple inventors from different CZs are assigned fractionally to each listed CZ.[4]

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. Additionally, we collect data on local employment by industry. Appendix B provides additional information regarding the data's construction.

Our unit of time throughout the analysis 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 number of patents 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. This provides a sample of 478 CZs, which accounted for approximately 92% of the U.S. population in 2010.[5]

[4] Berkes (2018) provides details about the data collection procedure, 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."

[5] This rule requires that cities had a population of at least 5,012 people in 1890 and 31,789 people in 2010.

Sectors are defined as the technological class-groups obtained by grouping 3-digit International Patent Classification (IPC) categories into 11 class-groups, as detailed in Appendix Table A.1.[6]

2.2 Historical trends

During the last 150 years, the technological landscape has undergone significant transformations, as measured by changes in the patenting shares of various patent classes. These changes are already 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 the patenting distribution since 1870.[7] The figure shows that 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.[8]

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.

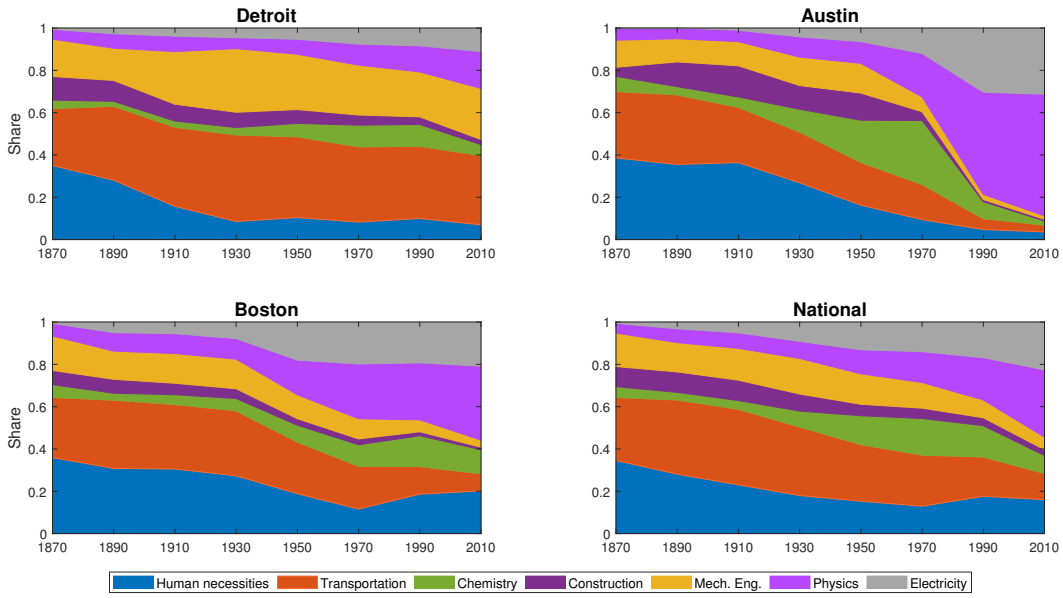
The remaining panels of Figure 1 depict three archetypal examples of this heterogeneity. Since the early 1900s, Detroit (top-left) has specialized in the production of patents related to "Transporting" and "Mechanical Engineering". In 1910, these two categories accounted for approximately 62% of its 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.

[6]Patents listing multiple three-digit IPC classes are assigned fractionally to class-groups, in proportion to the frequency of appearance of each class-group in the list of 3-digit IPC classes.

[7]For simplicity, 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.

[8]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**



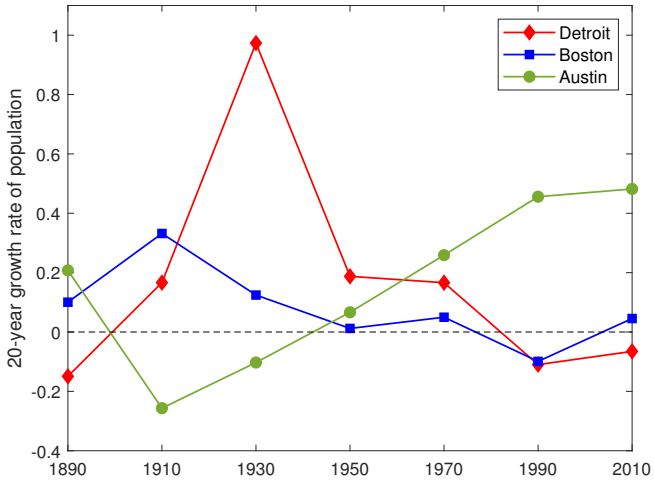
Notes: Composition of patenting output across the seven main IPC classes. Patent count for year t is constructed as the sum of patents filed between $t-10$ and $t+9$. Class names are abbreviated. The full description for each class is available at https://www.wipo.int/classifications/ipc/en/.

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 as measured by city patenting, 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, cities experience heterogeneous productivity gains from common technological shocks. Ultimately, this contributes 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

Figure 2: **City dynamics**



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

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. The diversification of Boston's patenting output may have made the city less sensitive to changes in the technological landscape, which could account for its stable growth trajectory.[9]

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. We now show that this pattern holds true over the long period covered by our data and is robust to controlling for potential local confounding variables such as innovation intensity, 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 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 setting, we construct the following measure of local exposure to technological

[9] Glaeser (2005) provides an overview of the causes of the slow decline of Boston between 1920 and 1980, and the subsequent re-emergence of the city. The high density of human capital is proposed as the major factor behind its resilience.

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 $m \neq n$ 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 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:

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

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.[10] $X_{n,t}$ denotes an additional set of controls that we discuss below.

The bin-scatter plot in Appendix Figure A.3 shows the relationship between the measure of exposure, $Exp_{n,t}$, and population growth, $\Delta \log (Pop_{n,t})$, residualized with respect to two lags of log-population and Census Division-time fixed effects. The graph reveals a strong positive correlation, indicating that cities with a more favorable initial exposure to the technological wave have experienced systematically greater population growth.

Table 1 displays the results of the regression. The estimates in column 2 (which controls for two lags of log-population and Census Division-time fixed effects) imply that an increase of one residual standard deviation in the measure of exposure is associated with an increase of 13.8% of one residual standard deviation in population growth. In column 3, we additionally control for lagged total patenting. This control does not significantly affect the exposure measure's coefficient, suggesting that the estimated relationship does not reflect a correlation between local innovation intensity and composition across fields. Thus, this correlation arises from comparing cities of

[10] 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).

similar size and 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 genuine scientific and technological advancements. Alternatively, they may be induced by 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. In other words, as emphasized in Schmookler (1966), our perspective is one of profit-driven innovation; the intensity of innovation across fields responds to changes in their returns, whether that be from changes in its costs (e.g., due to scientific advancements) or market size (e.g., due to changes in demand).

An alternative view to what we propose is one of innovation as a byproduct of production. If this is the case, then factors that differentially impact patenting across fields may be correlated with other industry-level shocks that drive differences in population growth across cities, which would confound our interpretation of the estimates. 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 the local shares and the aggregate shifts.[11] Appendix Figure A.4 plots this measure of exposure to industry shocks against our measure of exposure 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 relatively weak. The correlation is 0.138, and the R^2 of the underlying regression is 0.019. This suggests substantial differences in the composition of patents across fields that the local employment distribution across industries does not explain. 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 4 of Table 1, we estimate the relationship between exposure to technological waves and local population growth while directly controlling for industry shock exposure. This variable is a strong predictor of contemporaneous population growth, and its inclusion slightly reduces the estimated coefficient on the exposure to technological waves, which nevertheless remains large and significant at the 1% level.[12]

In column 5, 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 the measure of exposure to technological waves does not merely reflect the availability of human capital in the city.

[11] 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.

[12] Note that, if our previous interpretation is correct, this control would purge part of potentially valid variation.

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

	Growth rate of population				
	(1)	(2)	(3)	(4)	(5)
Exposure to tech. waves, $Exp_{n,t}$	0.637***	0.518***	0.453***	0.337***	0.327***
	(0.133)	(0.096)	(0.091)	(0.082)	(0.085)
Log-total patents			0.050***	0.015	0.011
			(0.014)	(0.012)	(0.012)
Industry composition				0.808***	0.807***
				(0.141)	(0.141)
Human capital (ranking)					0.037
					(0.051)
Log-population (lags 1 and 2)	Yes	Yes	Yes	Yes	Yes
Fixed effects	T	CD×T	CD×T	CD×T	CD×T
# Obs.	2,868	2,868	2,834	2,818	2,818
R^2	0.393	0.508	0.502	0.515	0.515

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

In conclusion, we have presented evidence in this subsection of 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 explains this relationship. We propose that this correlation can be partly explained by 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 subsection, 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.

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 15.7% of all citations for patents filed since 1940.[13] This proportion of citations is 6.32 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 4.36 times greater than the background probability and in the late subsample, it is 6.58 times greater.[14]

Analogously, we find substantial evidence of localization of patent citations in the technological space. Appendix Figure A.5 displays a heatmap of the probability that a citation from each technological 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 Model

This section develops a model incorporating an endogenous growth component into a spatial equilibrium framework. In our model, growth results from innovations that improve the current stock of ideas, and the diffusion of ideas across space and fields of knowledge is frictional. The theory formalizes the feedback between changes in the technological landscape and the evolution of the economic geography, and provides a rationale for the reduced-form relationship between population growth and local exposure to technological waves documented in Section 2.

We start by presenting the model in its simplest form. This version of the model allows us to solve for equilibrium in closed form, and we use it to derive the main theoretical predictions. In Section 4, we incorporate overlapping generations, costly migration, and local congestion forces into this model. This version is utilized to conduct the quantitative analysis of Section 5.

[13] 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.

[14] The observed probabilities are 14.0% in the early and 16.7% in the late sub-samples, whereas the background probabilities are 3.2% and 2.5%, respectively.

3.1 Environment

Consider an economy with 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.

3.1.1 Preferences, endowments, and demographics

In each period, a new generation is born in the location of their parents and makes decisions regarding migration and occupation. Individuals live for one period and have f_t children at the end of the period. There are no moving costs across locations.[15]

Agents make migration and occupational decisions to maximize expected utility, subject to idiosyncratic utility draws that affect each agent's desirability of each location-sector. Each individual $i \in L_t$ receives a complete set of stochastic utility draws at the beginning of the period, 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 $\zeta > 1$. If individual i chooses location-sector (n, s), they obtain utility

$$U_{n,s,t}(\mathbf{x}_i) = u_n \, x_{n,s,i} \, c_{n,s,i,t}, \tag{3}$$

where u_n is the level of amenities in city n and $c_{n,s,i,t}$ denotes consumption of the final good by individual i in location-sector (n, s) at time t.[16] Because their 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 3.2.2.

3.1.2 Production and innovation technologies

Each agent i is endowed with one unit of labor, which they supply inelastically with productivity q_i. Total output in the economy is given by a linear aggregator of individual productivity across all locations and sectors:

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

[15] We will relax these assumptions in Section 4 by allowing for overlapping generations and costly migration.

[16] The assumption of exogenous and time-invariant amenities will be relaxed in Section 4 by allowing for both an exogenous time-varying component and an endogenous component that captures local congestion forces.

where $L_{n,s,t}$ denotes the mass of agents in the location-sector (n,s) and $\mathbb{E}[q_{n,s,t}]$ denotes their average productivity.[17]

Individual productivity is endogenously determined by a decision to imitate or innovate. The quality of ideas in the choice set of each agent is stochastic, and its distribution varies by location-sectors.[18] At the beginning of each period, every agent i in the new generation receives 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, $z^l_{n,s,i}$, represents a random draw from the distribution of productivity among agents employed in location-sector (n,s) in the previous generation, whose cumulative distribution is denoted by $F_{n,s,t-1}(q)$. This draw can be interpreted as the knowledge that individual i acquires from their teacher, mentor, or manager, which can be imitated and directly applied to production.[19] If the agent adopts this idea, their lifetime productivity is

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

Imitators represent non-creative workers (such as those in routine occupations) who use previously developed ideas in the production process.

The second set of terms, $\{z^x_{m,r,i}\}_{m,r \in N \times S}$, represents a full vector of random draws from each productivity distribution in all locations and sectors in the previous generation, with corresponding cumulative distributions $\{F_{m,r,t-1}(q)\}_{m,r \in N \times S}$. Note that this complete set of draws contains local ones (i.e., $m = n$ and $r = s$). These draws can be interpreted as knowledge that the agent acquires through various transmission channels, such as books, radio, television, and the internet, or through casual interactions with local or non-local individuals. Although these ideas cannot be directly adopted in production, they can serve as an input for innovation. In particular, an agent employed in (n,s) can use an idea $z^x_{m,r,i}$ to innovate and achieve productivity

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

The term $d_{(m,r)\to(n,s)} \geq 1$ captures geographical and technological knowledge transmission frictions

[17] Equation (4) can be obtained as an equilibrium representation between aggregate output and total labor in a setting where the fundamental production function involves a constant returns to scale production function between labor and location-sector specific intermediates (see, for example, Chapter 14.1 in Acemoglu, 2009).

[18] Our description of knowledge diffusion builds on the model developed by Buera and Oberfield (2020), who study an environment in which the distribution of ideas endogenously converges to a Fréchet distribution.

[19] De la Croix et al. (2018) develop a model of knowledge diffusion in which the institutions controlling the effectiveness of knowledge transmission between journeymen and apprentices contribute to explaining differences across countries in long-run growth.

that diminish the efficacy of knowledge between the idea origin (m, r) and the idea destination (n, s). The term $\alpha_{r,t}$ represents the importance of sector r in the innovation landscape. The greater the value of $\alpha_{r,t}$, the more effectively knowledge in sector r can be developed into innovations for all sectors. For this reason, we refer to changes in $\alpha_{r,t}$ as technological wave shocks. Note that this term is unrelated to agent i's location-sector (n, s). Rather, it is a characteristic inherent to the sector of origin r at time t. Lastly, $\epsilon_{n,s,t}$ 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 productivity but are not accounted for in Equation (6), such as the opening of production and research facilities.

3.2 Equilibrium

We normalize each period's price of the final good to one. Thus, the wage of an 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:

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

3.2.1 Diffusion of knowledge

Agent i in location-sector (n, s) chooses whether to imitate or innovate to maximize their productivity given their vector of idiosyncratic idea draws $\mathbf{z}_{n,s,i}$:

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

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, combining Equation (7) with Assumption

A1, individual productivity at any time $t \geq 0$ is distributed Fréchet with shape parameter θ and with scale parameter evolving according to the following law of motion:[20]

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

Equation (9) is crucial to our theory, because it describes the dynamics of productivity distributions across location-sectors.[21] The scale parameter $\lambda_{n,s,t}$ governing the distribution of knowledge in the new generation 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, $\alpha_{r,t}$, as well as the local effectiveness of innovation, $\epsilon_{n,s,t}$, and is discounted by technological and physical distance between location-sector pairs, $d_{(m,r)\to(n,s)}$.

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 $\lambda_{n,s,t}$, local average productivity can be written as

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

where $\Gamma(\cdot)$ denotes the gamma function. In other words, we can easily compute the dynamics of the average productivity of each location-sector using Equation (9).

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 on an idea from any location-sector (m, r) at time t can be expressed as follows:

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

[20] Tractability can be maintained without assuming independence, as in Lind and Ramondo (2019).

[21] This result follows from the max-stability property, which states that the maximum of a vector of 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:

$$Pr(\max\{a_k\, z_k\}_{k=1}^{K} \leq y) = \prod_{k=1}^{K} Pr\left(z_k \leq \frac{y}{a_k}\right) = \prod_{k=1}^{K} \exp(-\lambda_k a_k^{\theta} y^{-\theta}) = \exp\left(-\sum_{k=1}^{K} \lambda_k a_k^{\theta} y^{-\theta}\right).$$

3.2.2 Migration and occupational choice

At the beginning of period t, agents in the new generation observe sectoral and local shocks (i.e., $\alpha_{r,t}$ and $\epsilon_{n,s,t}$ for all $n \in N$, and $r, s \in S$), but do not know their idiosyncratic idea draws. 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

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

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 corresponds to

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

3.2.3 Equilibrium Definition

We now have all the ingredients to define an equilibrium of the model.

Definition 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).

All equilibrium conditions listed in the preceding definition 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.

3.2.4 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).

From Equation (15), it is immediately apparent 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. 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^*.

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 demonstrates that the stationary value of the scale parameters 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. Conversely, Equation (16) can be interpreted as stating that, conditional on relative productivities, $\frac{\tilde{\lambda}^*_{m,r}}{\tilde{\lambda}^*_{n,s}}$, higher BGP growth can result from greater local and sectoral characteristics (α^*_r and $\epsilon^*_{n,s}$) or from reduced frictions to knowledge diffusion ($d_{(m,r)\to(n,s)}$).[22]

3.3 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 ($\alpha_{r,t} \equiv \alpha^*_r$). As a result, the relative size of cities along the BGP does not change. Next, we will characterize the model's dynamics after incorporating technological wave shocks.

We study the model's dynamics by log-linearizing the equilibrium conditions around the BGP. By log-linearizing our model, we can obtain sharp and intuitive characterizations of what drives city growth in response to technological waves. For our characterization of the log-linearized dynamics, we assume that the economy is in a BGP at time $t-1$ (i.e., the average productivity in each location-sector grows at the same rate, and as a result, the distribution of population across locations is constant). 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) 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 each location-sector (n, s) from the BGP, $\hat{\lambda}_{n,s,t}$, is equal to the sum over all sectors $r \in S$ of the sectoral*

[22] 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.

shock to r*,* $\hat{\alpha}_{r,t}$*, weighted by the reliance of innovation in* (n,s) *on ideas from sector* r*, as measured by the probability of building on ideas from sector* r *when innovating,* $\eta^*_{r\to(n,s)} \equiv \sum_{m\in N} \eta^*_{(m,r)\to(n,s)}$ *(defined in Equation 11):*

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

The implication of Proposition 2 is that, all else being equal, the *sensitivity* of local productivity $\hat{\lambda}_{n,s,t}$ to shocks to any given sector, $\hat{\alpha}_{r,t}$, is increasing in the probability of drawing knowledge from that sector to innovate, $\eta^*_{r\to(n,s)}$. The existence of frictions to knowledge diffusion implies that this reliance on ideas from sector r, $\eta^*_{r\to(n,s)}$, depends on how "close" sector r is to (n,s) in the geographical and technological space. Quantitatively, in light of the evidence discussed in Section 2.4 regarding the localized nature of knowledge diffusion, the local stock of knowledge in the same sector, $\lambda_{n,r}$, will play a decisive role in determining this reliance term. Furthermore, it is evident from Equation (13) that this stock of knowledge is closely related to the local share of the population employed in the same sector.[23] This implies that the sensitivity of local productivity to shocks in any given sector is strongly linked to the sector's prevalence in the local economy.

Consider now 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.[24] Equation (19) implies an intuitive condition for whether a city grows or shrinks relative to the rest of the economy. A location grows 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:

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

[23] To see this, note that in the limit case of $\theta = \zeta$, $\alpha^*_r = \alpha^*_s$, and $d_{(n,r)\to(n,s)} = \bar{d}$ for all $r,s \in S$, the reliance of (n,s) on ideas from r, $\eta^*_{r\to(n,s)}$, is equal to the local employment share of r in location n, $\pi^*_{r|n}$.

[24] 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 C. 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 (C.2), combined with Equation (C.3), 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 determining the spatial response to local shocks. A similar mechanism also emerges in Schubert (2021), who studies how migration responses controlled by pre-existing migration linkages make local shocks to house prices propagate to other cities.

Finally, we characterize changes in population shares as a function of BGP values and exogenous shocks only, rather than as a function of endogenous changes in the productivity parameters, $\hat{\lambda}_{n,s,t}$ (as derived in Equation 19). Combining Equations (18) and (19), we derive the following proposition, which summarizes the model's implied population dynamics in response to 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 a simplified version of the model 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 rest of the economy, that is, $\sum_{m\neq n} \pi^*_m \approx 1$, for all $n \in N$.*

Rewriting Equation (20) under Assumption A2, we obtain

$$\hat{\pi}_{n,t} \overset{A2}{\propto} b_t + \sum_{s\in S} \pi^*_{s|n}\hat{\alpha}_{s,t}, \tag{21}$$

where $b_t \equiv -\sum_{s\in S} \pi^*_{\cdot,s}\hat{\alpha}_{s,t}$, 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, cities' differential patterns of population growth depend on a weighted average of aggregate sectoral shocks, with the weights corresponding to the city's pre-existing pattern of specialization across sectors.[25] In other words, in the absence of knowledge flows across fields, the transmission of technological wave shocks to population dynamics is only determined by the local employment shares across sectors. Equation (21) thus provides a rationale for the shift-share functional form

[25] 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.

used to measure exposure to technological waves in the reduced-form analysis of Section 2.3.[26]

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

$$\hat{\pi}_{n,t} \overset{A2.2}{\propto} b_t + \sum_{s \in S} \sum_{r \in S} \pi^*_{s|n} \eta^*_{r \to (n,s)} \hat{\alpha}_{r,t}, \tag{23}$$

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, via 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 all cities are of negligible size, the rest of the economy is the same for all locations (as captured by the term b_t in Equation 23). When individual cities are of non-negligible size, the rest of the economy is location-specific, and Equation (20) takes this fact into account.

Taking stock and looking ahead To summarize, 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 translate into 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 innovation from different fields, as shown in Equation (23). In the following section, we will extend the model by introducing potentially relevant features from a quantitative standpoint (overlapping generations, costly migration, and local congestion forces) and then show how we bring the model to the data to quantitatively assess the importance of this mechanism.

[26] 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:

$$\hat{\pi}_{n,t} > 0 \overset{A2}{\iff} \sum_{s \in S} \pi^*_{s|n} \hat{\alpha}_{s,t} > \sum_{s \in S} \pi^*_{\cdot,s} \hat{\alpha}_{s,t}. \tag{22}$$

4 Quantitative model and calibration

The reminder of the paper is devoted to a quantitative analysis of the impact of technological waves on city dynamics. This section describes how we extend and calibrate the model. The following section reports our quantitative results and provides a counterfactual analysis.

4.1 Extended model for quantitative analysis

We extend the model from Section 3 along three dimensions. First, we relax the assumption that individuals live for a single period. Second, we introduce costly migration by assuming 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. A growing body of research contends that these channels are quantitatively significant (see, among others, Diamond, 2016 and Allen and Donaldson, 2020). This section shows that these channels can be incorporated into our framework without incurring a prohibitive loss of tractability. Appendix C extends the log-linearized dynamics around the BGP 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 make their migration and occupational decisions by selecting 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. Agent i in their young and old periods is endowed with one unit of inelastically-supplied labor and has productivity levels (q_i^y, q_i^o).

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:

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

where $u_{n,t}$ is the level of amenities in city n at time t, $\mathbf{x}_i$ is a vector of idiosyncratic utility draws 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:

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

where $\mathbb{E}[q^y_{n,s,t}]$ ($\mathbb{E}[q^o_{n,s,t}]$) 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 imitate or innovate process outlined in Section 3.2.1 to determine their youth productivity q^y_i 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 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.

Appendix C provides more details on the equilibrium conditions in this extended framework.

4.2 Overview of the calibration strategy

The model has a recursive structure that enables us to sequentially calibrate parameters and unobserved variables by making a small number of transparent assumptions on the mapping of model's objects to data on population, income, and patenting. As in the empirical analysis of Section 2, 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 for 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 end of our sample period.

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, and set the time-invariant parameter ζ externally. In the second step, we recover local amenities $u_{n,t}$, the path of local productivities $\lambda_{n,s,t}$, and aggregate fertility f_t, while simultaneously determining the time-invariant parameters θ, ω, and A^o. In the third step, we estimate the costs of knowledge transmission $d_{(m,r)\to(n,s)}$ 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 $\epsilon_{n,s,t}$.

4.3 First Step: Migration elasticity and gravity equation for migration flows

We assume migration costs between each pair of locations, $\mu_{m\to n}$, to be an exponential function of their geographical distance in kilometers, $D_{m,n}$:

$$\mu_{m\to n} = e^{\bar{\mu} D_{m,n}}. \tag{25}$$

We show in Appendix D that the probability that an individual born in m migrates to n, $\pi_{m\to n,t}$, takes the following gravity form:

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

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

We estimate Equation (26) via Poisson Pseudo Maximum Likelihood (PPML).[27] For the estimation of (26) and for the other time-invariant parameters pertaining to migration, population, and income, we utilize IPUMS data from 1990, for which the most recent and comprehensive demographic data are available. We obtain a statistically significant estimate of the composite semi-elasticity, $\zeta\bar{\mu}$, of 0.003 (see Appendix Table D.4). This estimate indicates that a 100-kilometer increase in distance reduces the migration probability by approximately 35%.

Migration elasticity (ζ) As explained in detail in Section 4.4, population is matched by construction, while average income is directly related to the local stock of patents. Therefore, the parameter ζ, that captures the elasticity of migration (or sectoral mobility) with respect to average income, cannot be identified separately. In our baseline, we set $\zeta = 6$, which is at the upper end of the range considered by Allen and Donaldson (2020), given that the elasticity of lifetime sectoral mobility is likely greater than the elasticity of migration. We explore robustness of our results to

[27] See Appendix D for details on the estimation.

this parameter choice in Appendix E.[28]

4.4 Second Step: Amenities and productivities

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

4.4.1 Productivity distribution

First, consider the scale parameters $\lambda_{n,s,t}$. These parameters 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 a higher local income, a greater ability to attract population, and a greater potential for future innovation. We draw on Schumpeterian endogenous growth theory to postulate (and later validate) a direct mapping between the stock of knowledge and the stock of patents in each location-sector. Then, we use our model structure to link the stock of knowledge in a location-sector to its scale parameter $\lambda_{n,s,t}$ and we use this mapping to recover the complete path of $\lambda_{n,s,t}$.

We establish a connection with Schumpeterian growth theory by interpreting each idea draw $q_{n,s,i,t}$ as a point in a quality ladder with $h_{n,s,i,t}$ steps, so that

$$\log(q_{n,s,i,t}) = a h_{n,s,i,t}, \tag{27}$$

where $a > 0$ is the step size of each quality improvement. We define the local stock of knowledge in location-sector (n, s) at time t as the average number of steps in the local productivity distribution of innovators, $\mathbb{E}[h_{n,s,t}]$. In other words, the knowledge stock captures the average number of steps climbed in the "knowledge ladder" by innovators. Using that ideas $q_{n,s,i,t}$ are distributed Frećhet with scale $\lambda_{n,s,t}$ and shape θ, we have that the stock of knowledge satisfies

$$\mathbb{E}[h_{n,s,t}] = \frac{\gamma + \log(\lambda_{n,s,t})}{a\theta}, \tag{28}$$

where γ denotes the Euler-Mascheroni constant.

We then assume the following parametric relationship between the stock of knowledge in location-sector (n, s) at time t and the corresponding stock of patents:

$$\mathbb{E}[h_{n,s,t}] = \log\left[\tilde{G}_t \times \left(1 + \sum_{\tau=0}^{\tau_{max}} \xi^\tau Pat_{n,s,t-\tau}\right)\right], \tag{29}$$

[28] 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.

where $Pat_{n,s,t}$ denotes the total number of patents filed at time t in location-sector (n, s), $0 < \xi < 1$, and $\tilde{G}_t$ is a time-varying factor.[29] In other words, we assume a concave (logarithmic) relationship between the accumulated flow of all patents in a location-sector discounted at a rate ξ, and the stock of knowledge in this location-sector. The time-varying factor $\tilde{G}_t$ captures changes in the innovative content of patents that are common to all location-sectors (e.g., patents becoming more defensive over time). Calibration of these parameters is discussed below.

Combining Equations (28) and (29), we obtain the following:

$$\lambda_{n,s,t} = G_t \times \left(1 + \sum_{\tau=0}^{\tau_{max}} \xi^{\tau} Pat_{n,s,t-\tau}\right)^{\sigma}, \tag{30}$$

where $\sigma \equiv a \times \theta$ represents the elasticity of $\lambda_{n,s,t}$ with respect to the observed stock of patents and $G_t \equiv e^{-\gamma} \times \tilde{G}_t^{\sigma}$. The elasticity σ converts the variation in the local stock of patents into meaningful variation in the average location-sector productivity. Thus, Equation (30) provides a mapping between the stock of patents in a location-sector at a given time (which we can observe) and a key component of our theory, the scale parameter of the Fréchet distribution in that location-sector, $\lambda_{n,s,t}$ (which is unobservable).[30]

We calibrate σ and θ to match the standard deviation of log-income per capita across cities and in the overall population in 1990, which are 0.19 and 0.67, respectively.[31] The constant G_t is set to produce an annual growth in per capita income of 2%.[32] We set $\xi = 0.5$ and $\tau_{max} = 2$, reflecting an assumption that the contribution of past patents to variation in the local stock of knowledge halves every 20 years and vanishes after 60 years. This depreciation rate is consistent with the 4% obsolescence in Caballero and Jaffe (1993).[33] In Appendix E, we show that we obtain similar results to our baseline when setting $\xi = 0.19$, which corresponds to the 8% obsolescence rate inferred in Anzoategui et al. (2019) from patent renewal rates. Lastly, we set $A^o = 2.11$, which implies a 1990 experience premium of 1.42 (Heathcote et al., 2010).

The mapping of $\lambda_{n,s,t}$ to the stock of patenting (Equation 30) incorporates a size effect in which, all else equal, larger cities have higher average productivity. The existence of a correlation between 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,

[29] We add one to the stock of patents to assign a meaningful value to cases in which patenting is zero.

[30] Equation (29) defines the stock of knowledge in a reduced-form way following the strategy in Hall et al. (2001). Alternatively, it is possible to define the probability of patenting as the probability of improving existing ideas, following the approach in Kortum (1997). This approach delivers a direct mapping between the scale parameter and the number of patents without relying on the reduced-form definition of the stock of knowledge, yielding an expression similar to Equation (30).

[31] The standard deviation of log-income in the overall population is taken from Krueger and Perri (2006).

[32] We choose units of the final good so that the geometric average of $\lambda_{n,s}^{\frac{1}{\sigma}}$ is equal to one in the first time period.

[33] This obsolescence rate is inferred from the rate of decline in patent citations.

local productivity spillovers, and higher availability of productive 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 population with respect to per capita income is empirically accurate. As we show in Section 4.7, the model closely replicates the empirical relationship between population and income despite the fact that this relationship is not directly targeted.

4.4.2 Local amenities and fertility

Consider now local amenities $u_{n,t}$. Given values for $\lambda_{n,s,t}$ and previously calibrated time-invariant parameters, we infer local amenities by exactly matching population by city for each model period between 1890 and 2010. To calibrate ω and thus separate the exogenous component of local amenities ($\bar{u}_{n,t}$) from its endogenous component ($L_{n,t}^{\omega}$), we impose as an identifying condition that the change in the exogenous component of local amenities should not be systematically correlated with the local exposure to technological waves ($Exp_{n,t}$), as defined in Section 2.3:

Assumption A3. *Let* $\widehat{\Delta \log \bar{u}_{n,t}}$ *and* $\widehat{Exp}_{n,t}$ *denote the values of* $\Delta \log \bar{u}_{n,t}$ *and* $Exp_{n,t}$ *residualized with respect to period fixed effects. Then:*

$$Cov\left(\widehat{\Delta \log \bar{u}_{n,t}}, \widehat{Exp}_{n,t}\right) = 0. \tag{31}$$

This condition enables us to identify ω, as lower values of ω imply a stronger correlation between changes in the exogenous component of amenities and local population growth. This restriction mandates that any systematic correlation between exposure to shocks and changes in local amenities must be attributable to the endogenous mechanism of the model and not the exogenous components. This condition is comparable to Ahlfeldt et al. (2015), who use the orthogonality between a neighborhood's distance from the Berlin Wall and the change in exogenous neighborhood characteristics to identify the strength of local agglomeration forces. We estimate $\omega = -0.092$, which falls within the range of values estimated by Fajgelbaum and Gaubert (2020) and Eckert and Peters (2022), among others, in different contexts.

Finally, we calibrate the fertility path, f_t, to match population totals by period. In Appendix Figure A.7, we show computationally that there are unique values of θ, σ, and ω that match their respective data moments.

4.5 Third Step: Gravity equation for knowledge flows

In the third step of calibration, we recover the parameters controlling knowledge transmission costs, $d_{(m,r)\to(n,s)}$. To this end, we derive a gravity representation for knowledge flows that we estimate

using data on patent citations. We parametrize knowledge diffusion frictions as multiplicatively separable between a geographical and a technological component:

$$d_{(m,r)\to(n,s)} = e^{\delta^G \mathbf{1}_{m\neq n} + \delta^K_{r\to s}}, \tag{32}$$

where δ^G controls the effectiveness of knowledge flows *across* locations relative to flows *within* locations, and $\delta^K_{r\to s}$ controls the applicability of ideas from sector r for innovation in sector s.

Combining Equations (11) and (32) and taking logs on both sides yields

$$\log(\eta^t_{(m,r)\to(n,s)}) = \phi^0_{m,r,t} + \phi^1_{n,s,t} - \theta\delta^G \mathbf{1}_{m\neq n} - \theta\delta^K_{r\to s}, \tag{33}$$

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

Equation (33) illustrates that bilateral citation probabilities $\eta^t_{(m,r)\to(n,s)}$ depend on the composite parameters $\theta\delta^G$ and $\theta\delta^K_{r\to s}$. We estimate Equation (33) using data on patent citations across location-sector pairs in order to recover those composite parameters. To obtain a more precise measurement of the flow of knowledge among inventors, we only include citations added by applicants. To achieve this, we limit 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 $\eta^t_{(m,r)\to(n,s)}$ as the proportion of citations from patents in (n,s) to patents in (m,r).[34]

We estimate this relationship using PPML and present the results in Appendix Table D.5. Using the estimated vaulue of the composite parameter $\theta\delta^G$ in conjunction with the value of θ, we obtain $\delta^G = 3.63$. The coefficient indicates that knowledge flows are highly localized, with the effectiveness of transmission across locations, $e^{-\delta^G}$, estimated to be approximately 2.66% of the effectiveness of transmission within locations. Notably, despite this low value, 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 $m \neq n$.

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 D.13. 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").

[34]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. In other words, 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 for half and identical to the local distribution of patents across technology classes for the remaining half. This also ensures that all patents, even those without backward citations, are included in the estimation. Appendix D provides additional details on the estimation.

Table 2 summarizes values and targets of the parameters calibrated up to this point.

4.6 Fourth Step: Technological waves and structural residuals

In the fourth calibration step, we combine the previously estimated values with the law of motion of 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 $\{\alpha_{s,t}\}_{s\in S}$. To complete the identification, we need to impose S additional conditions. We proceed with the following identifying assumption: 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.

To implement our identification, we make the following assumption:

Assumption A4. *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:*

$$\mathbb{E}\left[\epsilon^{\theta}_{\cdot,s,t}\right] = 1, \qquad \forall s \in S,\, t \geq 0. \tag{34}$$

Combining Equation (34) 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 $\{\epsilon_{n,s,t}\}$. In practice, based on the mapping we established in the first step between knowledge stocks and scale parameters, 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.

Thus, our 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 will pick up as a technological wave any expected common sectoral shock that redirects innovation across industries.

Appendix Figure A.10 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",

Table 2: **Parameter values and targets**

Parameter	Value	Target	Model	Data
		Calibrated parameters		
σ	0.215	s.d. log-income (across CZs), 1990	0.191	0.191
θ	2.10	s.d. log-income (overall), 1990	0.67	0.67
A^o	2.11	Experience premium, 1990	1.42	1.42
ω	-0.092	Orthogonality Condition (31)	0.0	0.0
		Assigned parameters and gravity equations		
ζ	6.0	Allen and Donaldson (2020)		
$\bar{\mu}$	0.0005	Table D.4		
δ^G	3.63	Table D.5		
$\delta^K_{r\to s}$	Figure D.13	Table D.5		

Notes: Standard deviation of log-income for the overall population is taken from Krueger and Perri (2006). Experience premium is taken from Heathcote et al. (2010). Standard deviation of log-income across CZs are author's calculations from the NHGIS.

and "Chemistry", and significant expansions in "Health" and "Physics", with the latter containing the majority of IT and computing-related innovation.

4.7 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. This allows us to illustrate and quantify the underlying mechanisms in a transparent way. We now show that, despite its parsimony, the 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 Appendix Figure A.8 depict the relationship between log-population and log-income in the model and in the data. Since the earliest model period for which per capita income is observed in publicly available data at the CZ-level is 1990, we show this relationship in 1990 (left-panel) and 2010 (right-panel). The model generates an elasticity of population with respect to per capita income which is empirically accurate, despite the fact that this elasticity is not targeted by the calibration.[35] The scatter plots in Appendix Figure A.9 display the relationship between log-income in the model and in the data across all cities in 1990 (left-panel) and 2010 (right-panel). 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

[35] In 1990 (2010), the slope of the regression line is 0.138 (0.118) in the model and 0.111 (0.082) in the data.

underlying regressions are equal to 0.62 and 0.45, respectively.

5 Quantitative results

In this section, we use the calibrated model to assess the role of frictional idea diffusion on local growth. Our main finding is that both geographical and technological frictions play a substantial role in explaining the empirical relationship between exposure to technological waves and local growth in the United States during the twentieth century. Then, we discuss two applications of our quantitative framework. First, we investigate the relationship between local specialization and growth volatility. We find that frictions to idea diffusion make local growth in specialized cities more volatile than in diversified cities. Second, we use our framework to explore how future technological trends may reshape the economic geography of the United States in the coming decades, revealing large geographical effects of future scenarios.

5.1 Technological waves, local growth, and knowledge diffusion

The primary objective of our analysis 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).[36] 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 aggregate ($\alpha_{s,t}$), and local shocks ($\epsilon_{n,s,t}$ and $\bar{u}_{n,t}$). By construction, this version of the model reproduces the data exactly. Thus, the resulting relationship between exposure to technological waves and local population growth is identical to the empirical relationship estimated in Section 2.3. Column 1 of Table 3 reports the coefficient obtained by regressing local population growth on exposure to technological waves with the full set of controls (Equation 2), which corresponds to the estimated coefficient in column 5 of Table 1.

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. 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

[36] In the absence of shocks, this assumption implies that city size would remain constant indefinitely.

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 57.1% lower than the empirical one. This suggests that the model's mechanism of frictional idea diffusion explains more than half 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.[37]

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, $\eta^t_{s\to(n,s)} \approx 1$ (Assumption A2.1). Second, frictions to idea diffusion *across locations* imply that cities with a greater concentration of expanding fields will experience greater productivity growth in *all* sectors due to localized knowledge spillovers between 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 $r \neq s$).[38] We then compute the model forward from the initial 1890 BGP until the last period (2010) under the same 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 36.5% less than in the complete model of column 1. Compared to the estimate in column 2, this number suggests

[37]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 adoption 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.

[38]Incidentally, this formulation is equivalent to assuming that diffusion across locations is *frictionless* (i.e., $\delta^G = 0$), since, in both cases, the reliance of local innovators on ideas from any given sector, $\eta^t_{r\to(n,s)}$, 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.

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 ($Exp_{n,t}$)	0.327***	0.140	0.207**
	(0.085)	(0.085)	(0.084)
Difference from empirical coefficient	-	0.188	0.121
Share explained by tech waves ($\alpha_{s,t}$)		57.1%	36.5%
Decomposition:			
- Share explained by tech. frictions			64.0%
- Share explained by geo. frictions			36.0%
Log-population (lags 1 and 2)	Yes	Yes	Yes
Other controls	Yes	Yes	Yes
Fixed effects	CD×T	CD×T	CD×T
# Obs.	2,818	2,818	2,818
R^2	0.52	0.50	0.51

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 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. $^{**}p < 0.05$, $^{***}p < 0.01$.

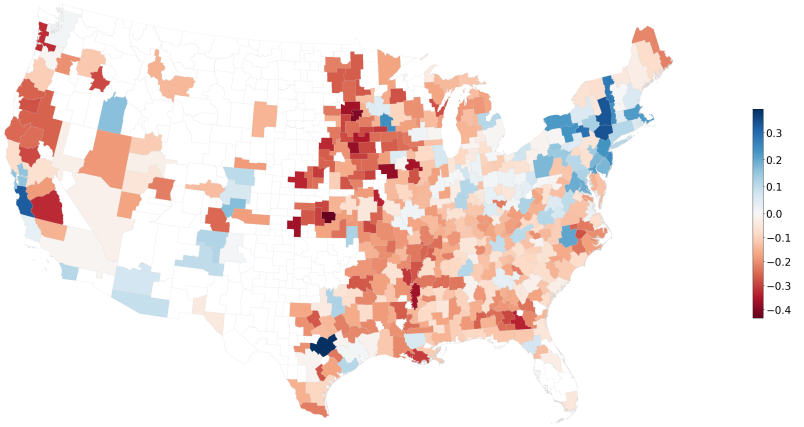
that approximately 64% of the overall mechanism can be attributed to frictions to diffusion across technological fields, while the remaining portion (approximately 36%) can be attributed to frictions to diffusion across locations. In other words, localized knowledge flows across fields of knowledge amplify the direct impact of sectoral shocks significantly.

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

We now use our framework to evaluate the extent to which technological waves coupled with knowledge diffusion frictions have contributed, throughout the last century, to delineating the current economic geography of the United States.

Figure 3 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

Figure 3: **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.

2 of Table 3). Darker colors represent larger differences between the actual and counterfactual growth rates, indicating cities that were more (or less) favorably affected by the technological waves that unfolded since the late 1800s.

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 32.4%, 38.8%, and 20.3% smaller, respectively. Interestingly, Seattle—another major technological hub—does not appear to have been significantly affected. The reason 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 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. For example, in the absence of technological waves since 1890, population in Omaha, Lincoln, and Des Moines would have been 16.0%, 10.8%, 24.5% larger, respectively.

5.2 Specialization and volatility of local growth

The two channels that control local population dynamics in response to technological wave shocks, reflecting frictions to idea diffusion across fields and across locations, make growth in diversified cities less volatile than in specialized ones. First, frictions to idea diffusion *across fields* imply that

productivity growth in a particular sector is primarily driven by shocks within that sector. Second, frictions to idea diffusion *across locations* imply that the local reliance on ideas from a particular field is increasing in the local availability of ideas from that field. Consequently, diversified cities, whose sectoral composition is dispersed across multiple sectors, and whose innovators rely on ideas from a broader range of fields, experience a less volatile trajectory in productivity and ultimately in population.

Exploring this link formally requires us to define the correct local specialization measure. Appendix F shows that under intuitive conditions, the variance of local population growth equals the Euclidean distance between the local and national vectors of sectoral shares. We use this distance as our measure of local specialization. Formally,

$$Spec_{n,t} \equiv \sum_{s \in S} \left(\pi_{s|n,t} - \pi_{\cdot,s,t} \right)^2, \tag{35}$$

where $\pi_{\cdot,s,t}$ is the share of the national population employed in sector s. According to this measure, cities are perfectly diversified if their local sectoral shares are exactly equal to the national shares.

To quantify the effect of diversification on the volatility of city growth, we conduct 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., over a forty-year horizon, or until 2050). These shocks are drawn from a normal distribution with mean zero and standard deviation equal to the standard deviation of the calibrated $g^{\alpha}_{s,t}$ throughout the sample. Then, we compute the counterfactual equilibrium for each simulation and correlate the standard deviation of population growth across simulations with 2010 specialization (as defined in Equation 35).

Column 1 of Appendix Table A.3 reports the coefficient of a regression of the standard deviation of local growth across the 5,000 simulations on the specialization measure and 2010 log-population (to account for size effects on specialization and volatility). As predicted by the theory, the volatility of population growth of specialized cities is significantly higher across simulations. The magnitude of the coefficients implies that the standard deviation for cities at the 90th percentile of the specialization distribution is 3.1 percentage points greater than for cities at the 10th percentile (the average volatility is 4.2 percentage points). The standard deviation increases by 5.6 percentage points as specialization increases from the level of Boston (a highly diversified city) to the level of Austin (a highly specialized one).

To disentangle the role of frictions across fields and across locations in explaining this relationship, we conduct the same experiment restricting knowledge diffusion to within-fields flows only (i.e., by imposing $\delta^{K}_{r \to s} = \infty$ if $r \neq s$). Our results are reported in column 2 of Appendix Table A.3. The coefficient for the measure of specialization decreases by 10.6% 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 technological waves on the U.S. geography

The quantitative model can be used to study how the economic geography will change in the coming decades in response to technological transformations. In this section, we examine which cities will be most positively and negatively affected by possible scenarios for future technological waves. Specifically, we project population growth across cities until 2050 under various hypotheses on the evolution of the importance of different sectors ($\alpha_{s,t}$), and compare the result to a baseline scenario in which the importance of all sectors is held constant at their 2010 levels.

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.[39] This scenario is conceivable if new advancements in transit technologies and autonomous vehicles cause transportation innovation to resume its central role. The left map in Appendix Figure A.11 illustrates the results. Compared to the baseline, CZs in blue (red) experience a net gain (loss) of population. Rust Belt cities are the areas best positioned to benefit from this transformation. According to our experiment, Detroit's population would increase by 3.8% relative to the baseline. Other centers of manufacturing would also benefit to a lesser extent. For instance, Gary would experience a relative population increase of 1.7%. A relative loss of population would be experienced by the three knowledge hubs of Austin (-1.9%), San Jose (-1.6%), and Seattle (-0.3%).

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 strengthening interdependence as a reduction in the cost of knowledge transmission ($\delta_{s \to r}$) by assuming a 50% decrease in composite knowledge frictions ($d^{\theta}_{(m,r)\to(n,s)}$) 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 the right map of Appendix Figure A.11. Also in this scenario, Detroit experiences a population increase (+2.3%), while San Jose experiences a net population loss (-0.5%). 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. Seattle experiences a relative population increase (+0.6%) due to its more diversified portfolio in IT and transportation.

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 pharmaceut-

[39] A 20% shock corresponds to approximately 2.5 times the standard deviation of calibrated shocks. This choice only affects the magnitude of the results but does not alter the qualitative patterns.

icals and medical sciences) possibly in response to new global health challenges. The results are depicted in the left map of Appendix Figure A.12. The experiment suggests that major CZs in the Northeast, such as Boston (+3.9%) and Providence (+8.7%), and in California, such as Los Angeles (+1.1%) and San Francisco-Oakland (+0.7%), would experience a net inflow of population at the expense of IT clusters such as Austin (-3.5%), San Jose (-3.2%), and Seattle (-1.3%).

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 the right map of Appendix Figure A.12. Under this scenario, population growth shifts from the East and West coasts and Rust Belt to the Central States. Des Moines (IA) has the greatest net gain (+17.4%) among the major CZs. This scenario would result in a population convergence across regions. A regression of log-population in 2010 on the predicted growth rate between 2010 and 2050 yields a coefficient of -1.2%, indicating that the population would move away from larger CZs and toward less-populated areas.

6 Conclusions

The economic geography of countries is characterized by rich and heterogeneous dynamics. Some cities remain prosperous over long time periods, whereas others experience periods of rapid expansion and contraction. In this paper, we explore the hypothesis that these rich dynamics are driven in part by cities' patterns of specialization across fields of knowledge, 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 framework that combines elements from quantitative spatial models and theories of endogenous growth via innovation and idea diffusion. The model is highly tractable and it provides a wide range of predictions regarding how the economic geography of countries responds to changes in the technological environment.

The quantitative results suggest that our proposed mechanism of frictional knowledge diffusion can account for more than half of the reduced-form correlation between local population growth and exposure to technological waves. Approximately one-third and two-thirds of this effect is attributable to frictions to diffusion across geographical areas and fields of knowledge, respectively. The effects of technological waves are heterogeneous across cities and have contributed to delineating the current U.S. economic geography. The channels emphasized in the model imply that specialized cities have significantly more volatile growth trajectories than diversified cities. Counterfactual experiments suggest that future technological scenarios may have large and het-

erogeneous 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 tangled. This implication is critical to our understanding of how changes in technology have led, throughout the last century, to today's spatial distribution of activities, and how we should expect this distribution to change as the technological environment will evolve in future decades.

Our results can be used to inform the design of local and place-based interventions. Policy efforts aimed at attracting new firms and shaping the local industrial 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 of insurance and redistributive mechanisms in the design of local and national development policies.

References

Acemoglu, D. (2009): *Introduction to modern economic growth*, Princeton, NJ [u.a.]: Princeton Univ. Press.

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

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 (2020): "Persistence and path dependence in the spatial economy," Tech. rep., National Bureau of Economic Research.

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

Anzoategui, D., D. Comin, M. Gertler, and J. Martinez (2019): "Endogenous Technology Adoption and R&D as Sources of Business Cycle Persistence," *American Economic Journal: Macroeconomics*, 11, 67–110.

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.

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

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.

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, S., L. Caliendo, F. Parro, and W. Xiang (2022): "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.

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.

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., 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.

Hall, B., A. Jaffe, and M. Trajtenberg (2001): "The NBER patent citation data file: Lessons, insights and methodological tools," *NBER working paper 8498*.

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.

Krueger, D. and F. Perri (2006): "Does income inequality lead to consumption inequality? Evidence and theory," *The Review of Economic Studies*, 73, 163–193.

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

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

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.

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

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.

Schmookler, J. (1966): *Invention and Economic Growth*, Cambridge: Harvard University Press.

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.

Online Appendix
(for online publication only)

Technological Waves and Local Growth

by Enrico Berkes, Ruben Gaetani, and Martí Mestieri

June 2023

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	3,824	294,896	810,284	54	1.79e+07
Log-population	3,824	11.63	1.30	3.99	16.70
Population growth	2,868	.243	.370	-.543	3.59
Total patents	3,346	1,114	4,775	0	77959
Exposure to tech. waves	2,868	-.108	.090	-.446	.403
Industrial composition	2,847	-.153	.159	-.688	.250

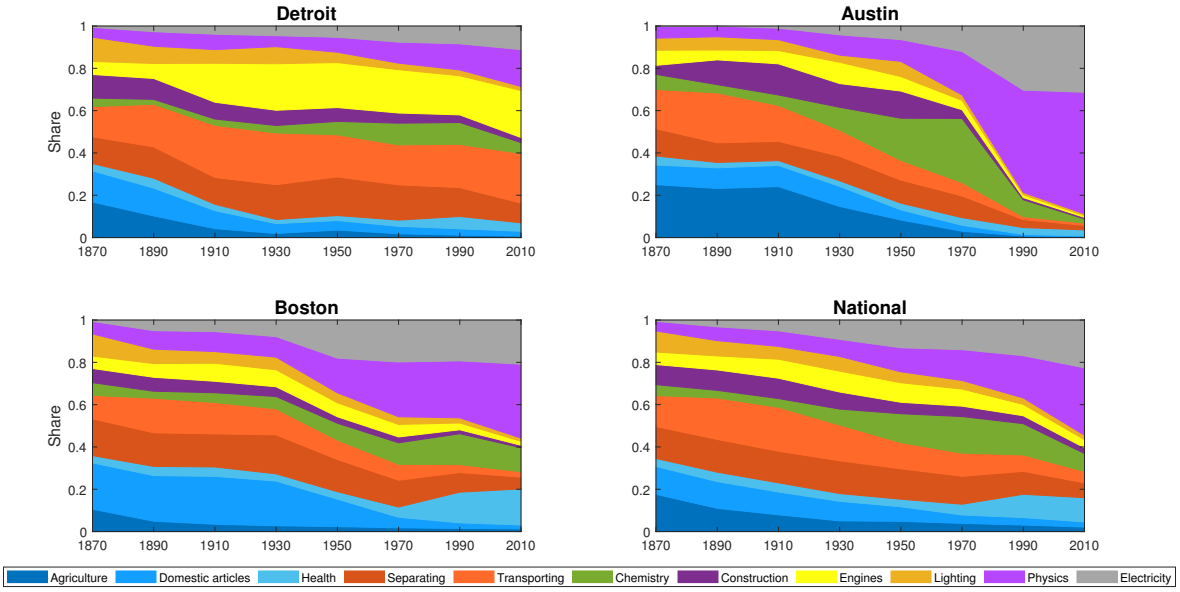
Notes: Summary statistics refer to the period 1870-2010 (Population, log-population), 1890-2010 (total patents), or 1910-2010 (remaining variables).

Table A.3: **Specialization and volatility of population growth**

	Dependent var.: Standard deviation across simulations	
	(1)	(2)
Specialization in 2010 ($Spec_{n,2010}$)	0.854***	0.764***
	(0.028)	(0.035)
Log-population in 2010	-0.002***	-0.004***
	(0.0004)	(0.0005)
Diffusion across fields	Yes	No
$Spec_{n,2010}$, 90th-10th perc.	0.037	0.037
Mean of dependent var.	0.042	0.044
Fixed effects	CD	CD
# Obs.	478	478
R^2	0.71	0.59

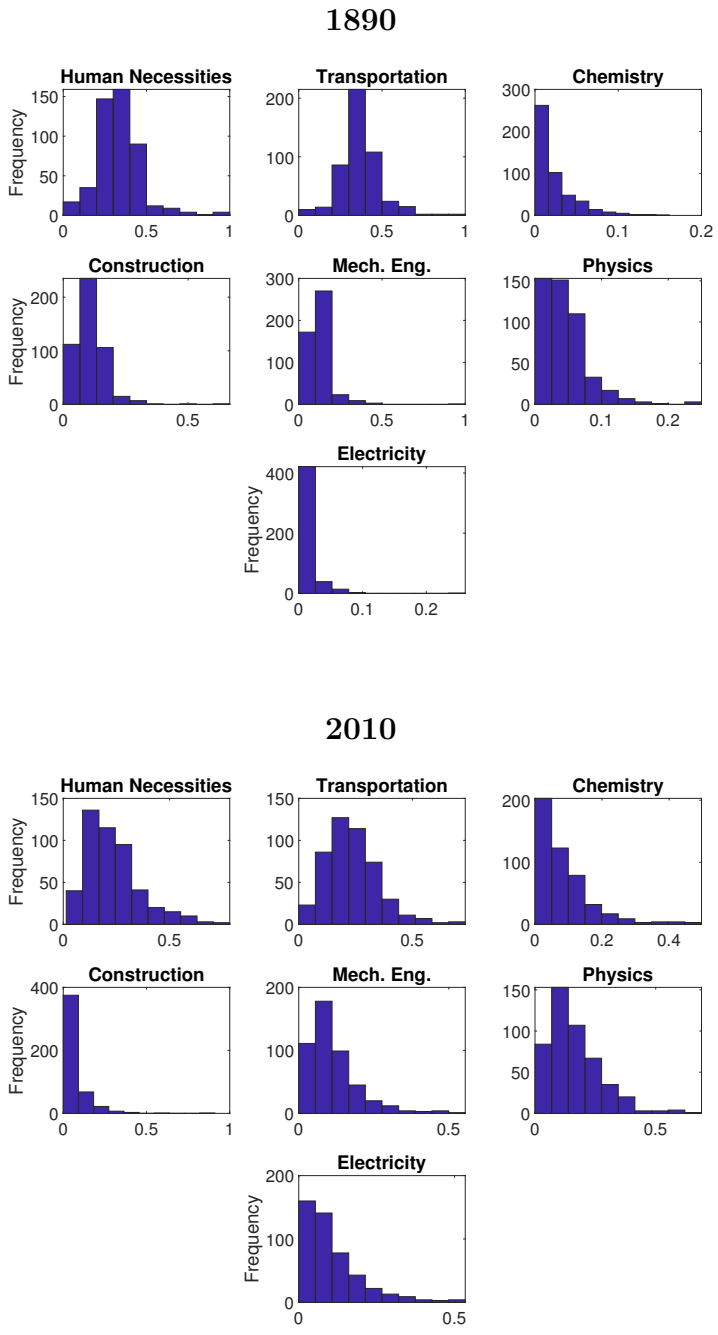
Notes: OLS estimates. Specialization is defined as in Equation (35). The dependent variable is defined as the city-level standard deviation of population growth across 5,000 simulations. ***$p < 0.01$.

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



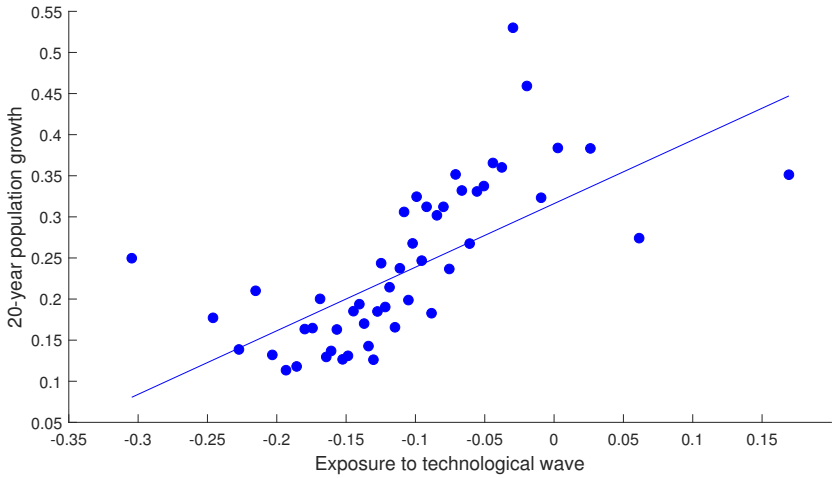
Notes: Composition of patenting output across the 11 IPC class-groups in Appendix Table A.1. Patent count for year t is constructed as the sum of patents filed between $t-10$ and $t+9$. Class names are abbreviated. The full description of each class is available at https://www.wipo.int/classifications/ipc/en/.

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



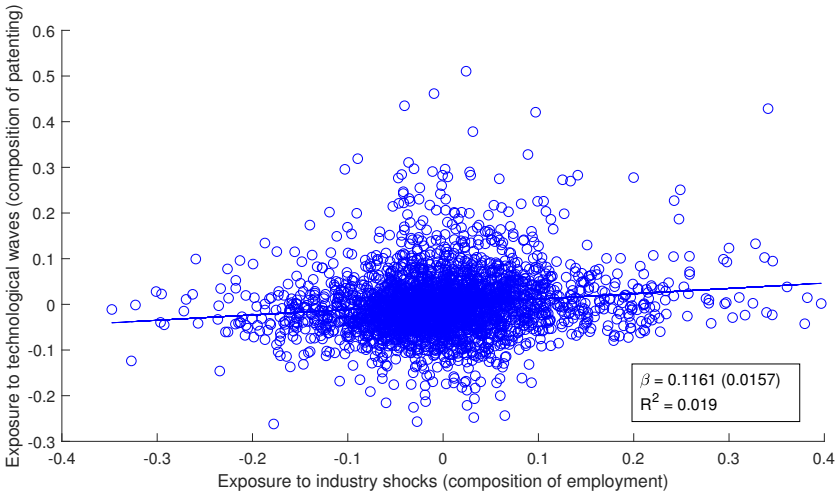
Notes: Distribution of patenting shares across cities for the seven main IPC classes in 1890 (top panel) and 2010 (bottom panel). Patent count for year t is constructed as the sum of patents filed between $t-10$ and $t+9$. Class names are abbreviated. The full description for each class is available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

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



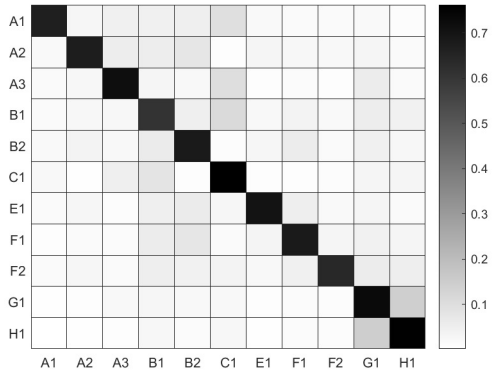
Notes: Bin-scatter plot of exposure to the technological wave, as defined in Equation (1), and 20-year population growth, 1910-2010. The bin-scatter plot is residualized with respect to Census Division-time fixed effects and two lags of log-population.

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



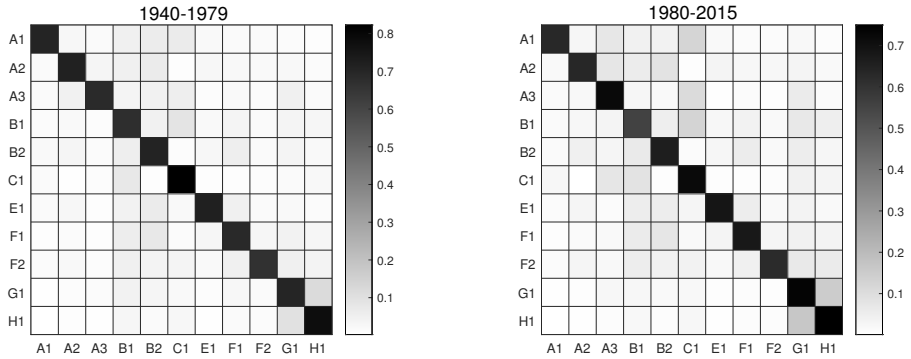
Notes: Scatter plot of exposure to the technological wave, as defined in Equation (1), and exposure to industry shocks, constructed using employment by industry, using the 12 main industries in the 1950 Census Bureau industrial classification system. Both variables are residualized with respect to two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

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



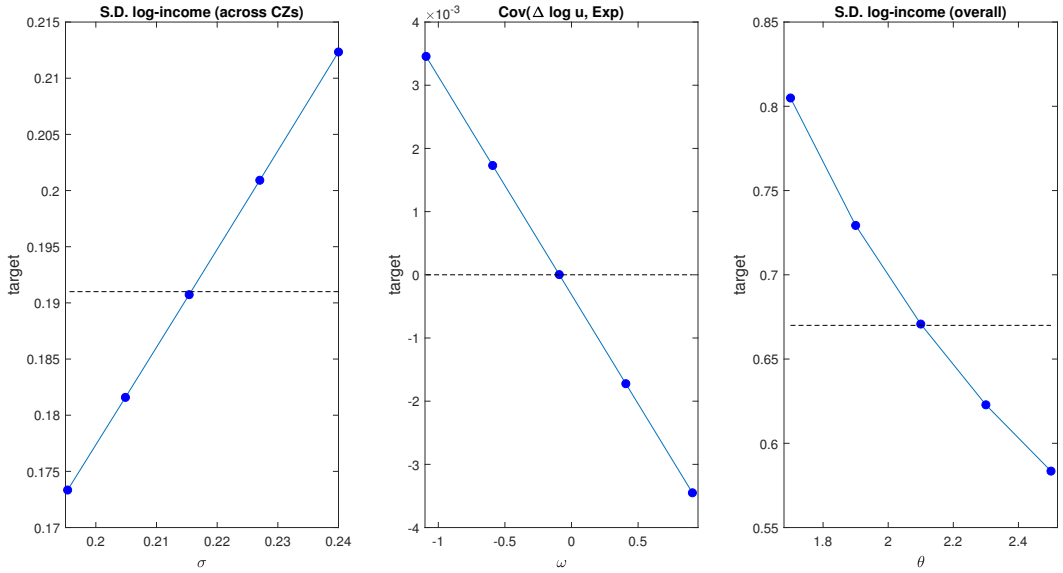
Notes: Probability that patents from the class-group on the vertical axis cite patents from the class-group on the horizontal axis. Probabilities are computed using patents filed since 1940 in which at least one inventor is in one of the 478 commuting zones in the main sample. Each citation is weighted 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**



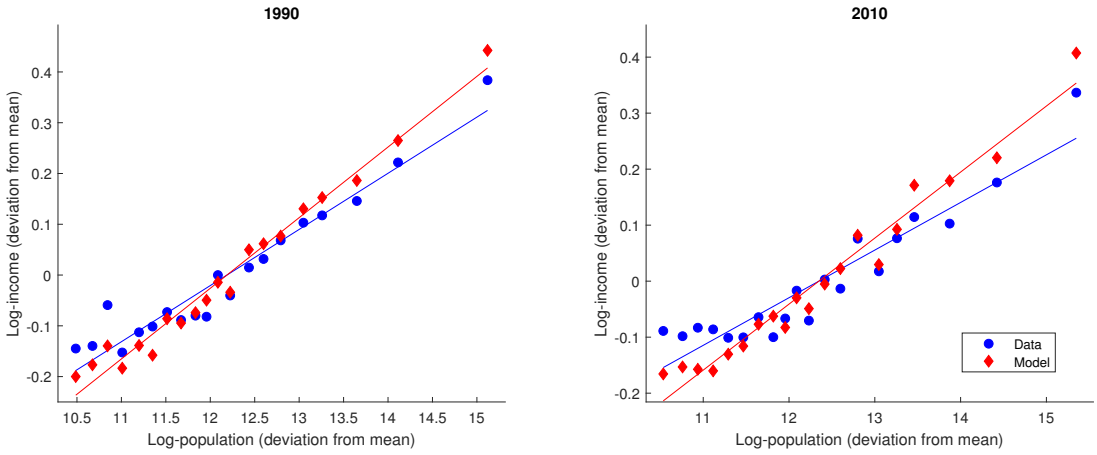
Notes: Probability that patents from the class-group on the vertical axis cite patents from the class-group on the horizontal axis. Probabilities are computed using patents filed since 1940 in which at least one inventor is in one of the 478 commuting zones in the main sample. Each citing patent is assigned a total weight of one. Class groups are described in Appendix Table A.1.

Figure A.7: **Identification of σ, ω, and θ**



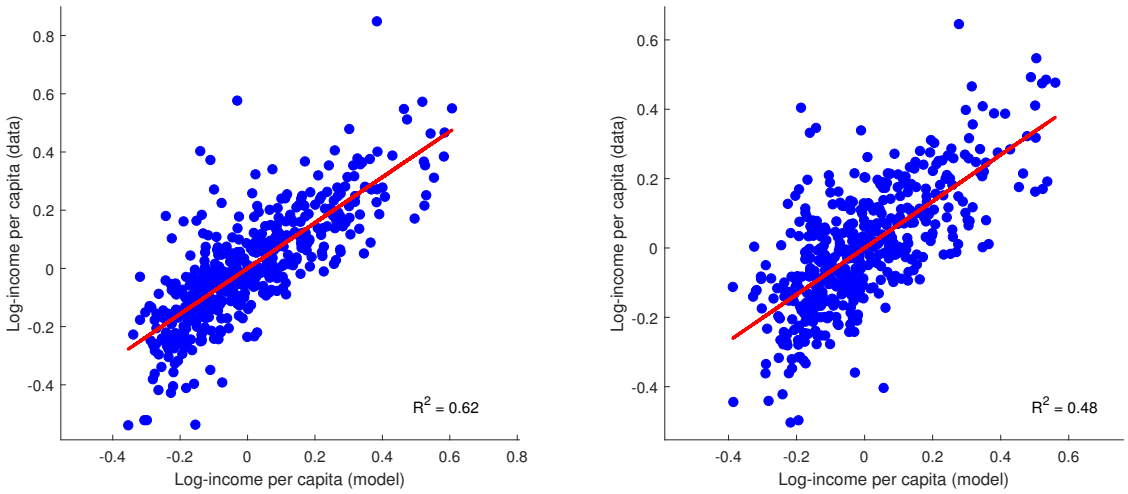
Notes: Moments in the data (horizontal dotted line) and in the model (blue marked line). Each of the plots is obtained by keeping the other two parameters constant at their calibrated values.

Figure A.8: **Population and income per capita: data and model**



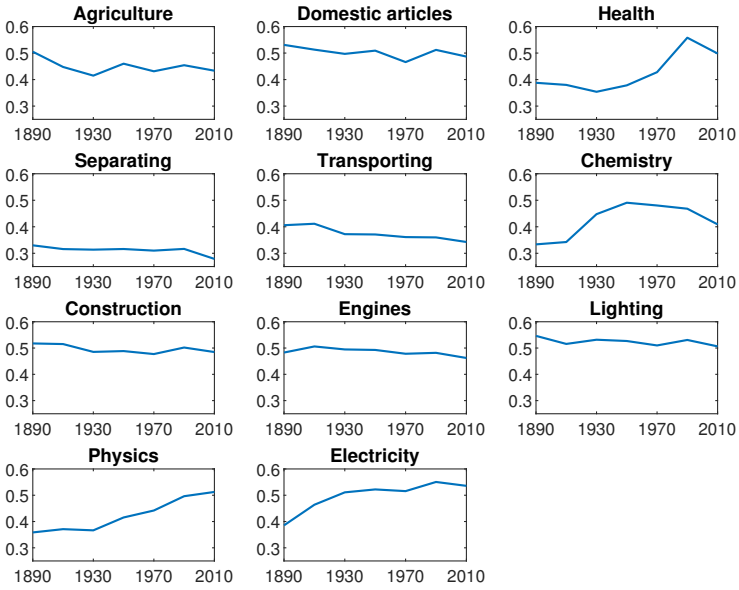
Notes: Bin-scatter plots of the relationship between log-population and log-income per capita in the data (blue) and the model (red) in 1990 (left-panel) and 2010 (right-panel). Log-income is displayed as deviation from the mean.

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



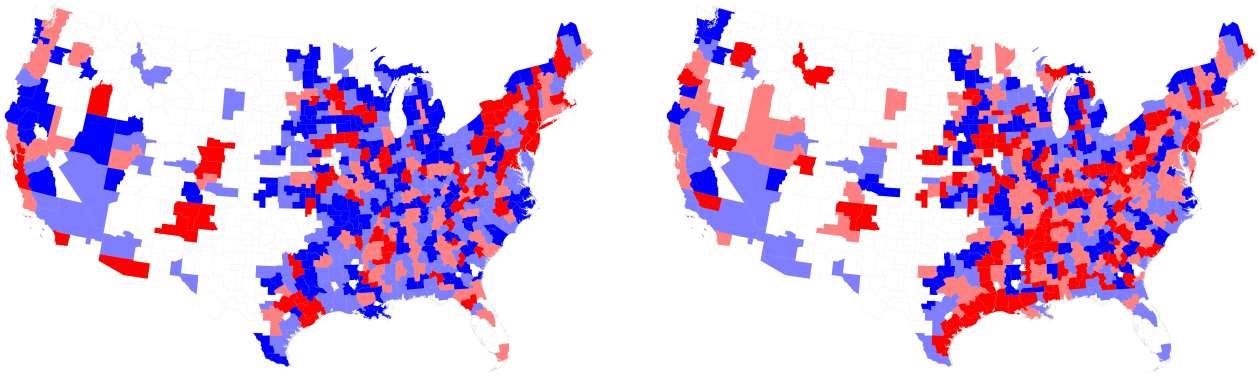
Notes: Scatter plots of log-income per capita in the model (horizontal axis) and data (vertical axis) in 1990 (left-panel) and 2010 (right-panel). All variables are displayed as deviations from the mean.

Figure A.10: **Calibrated technological wave shocks ($\alpha_{s,t}$)**



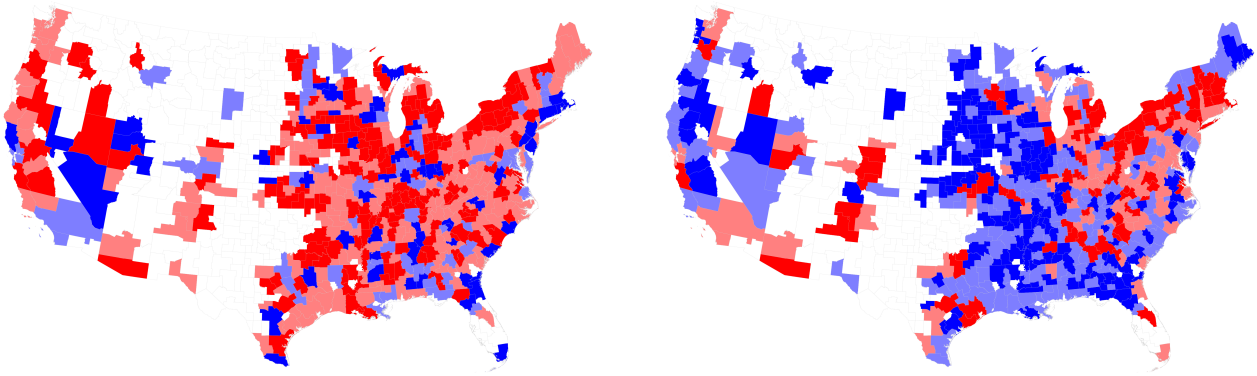
Notes: Paths of calibrated technological wave shocks ($\alpha_{s,t}$), 1890-2010. Class names are abbreviated. The full description of each class is available at https://www.wipo.int/classifications/ipc/en/.

Figure A.11: **Future scenarios: Autonomous vehicles**



Notes: The maps show log-population in 2050 after technological wave shocks of magnitude +20% to B2 (left map), as well as a 50% decline in composite knowledge frictions ($d^{\theta}_{(m,r)\to(n,s)}$) from (to) B2 to (from) both G1 and H1 (right map), 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.12: **Future scenarios: Pharmaceuticals and Agriculture**



Notes: The maps show log-population in 2050 after technological wave shocks of magnitude +20% to A3 (left map) and to A1 (right map) 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.[40]

We use the full-count censuses until 1940 to build crosswalks from historical counties to 1990 commuting zones. We then use these crosswalks in combination with county-level data from the NHGIS to consistently assign population and human capital data at the commuting zone level. Specifically, we follow a three-step procedure. First, we attempt to assign to each unique location in the historical decennial censuses – in terms of town, county, and state – their latitude and longitude.[41] Second, we count the number of people living in each town for the subset of locations that we were able to geolocate in the previous step, and reweight each town in this sub-sample so that the overall population count matches the county level data from the NHGIS.[42] Third, we assign the resulting town population to 1990 commuting zones.[43] For the decades 1950 onwards, we build crosswalks from counties to 1990 commuting zones based on the shares of overlapping areas. 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, ranging between 0.0006% of the U.S. population in 2010 and 0.02% in 1970.

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

[40] 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).

[41] We retrieve the coordinates from Google Maps or, when uniquely available, from an offline database available at https://nationalmap.gov.

[42] Some towns in newly annexed territories are occasionally reported with generic names such as *Township 43*. We drop these observations from the sample and reweight the remaining ones to match the county level population data.

[43] As an example, consider the town of Denver, CO, that in 1890 was part of Arapahoe County, a large and sparsely populated county. By 1990, the city of Denver had separated from the rest of the county to form its own. The first two steps reveal that a large portion of Arapahoe County's population in 1890 was located in Denver. The third step uses this information to correctly assign the largest share of population to the city and county of Denver. There are two special cases that are worth mentioning. First, when more than 95% of the area of the historical county falls within a 1990 county, then we assign the whole population to the 1990 commuting zone that contains the 1990 county. Second, when a historical county does not contain any town that we were able to geolocate reliably, then its population is assigned to 1990 commuting zones based on the overlapping of their areas.

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).

C 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. Agent i in their young and old periods is endowed with one unit of inelastically-supplied labor and has productivity levels (q_i^y, q_i^o).

Denoting by $L_{n,s,t}^y$ and $L_{n,s,t}^o$ the mass of young and old agents, the following identity holds:

$$L_{n,s,t}^o \equiv L_{n,s,t-1}^y.$$

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

$$L_{n,t}^k \equiv f_t \sum_{s \in S} L_{n,s,t}^y.$$

Migration and occupational choices are made to maximize expected lifetime utility, subject to migration costs and idiosyncratic utility draws. Utility of individual i born in location m and choosing location-sector (n, s) is given by

$$U_{m \to (n,s),t}(\mathbf{x}_i) = u_{n,t} \frac{x_{n,s,i} \left(c_{n,s,i,t}^y\right)^{\beta} \left(c_{n,s,i,t+1}^o\right)^{1-\beta}}{\mu_{m \to n}},$$

where $u_{n,t}$ is the level of amenities in city n at time t, $\mathbf{x}_i$ is a vector of idiosyncratic utility draws from a Fréchet distribution with shape parameter $\zeta > 1$, $\mu_{m \to n}$ represents migration costs of moving from m to n, $c_{n,s,i,t}^y$ and $c_{n,s,i,t+1}^o$ 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 across time and generations. Hence, individual consumption is equal to individual production in every period.

Young and old adults produce the final good using their unit of time according to their idiosyncratic productivity q. Thus, total output in the economy is given by a linear aggregator over

individual productivity across all locations and sectors:

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

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

Upon moving, young adults go through the imitate or innovate process described in Section 3.2.1 to determine their youth productivity q^y. Note that the relevant variables ($\epsilon_{n,s,t}$ and $\alpha_{n,s,t}$) are known at the time of the migration and occupational decision. 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

$$\pi_{m \to (n,s),t} = \frac{\left(u_{n,t} \frac{\lambda_{n,s,t}}{\mu_{m \to n}} \right)^{\zeta}}{\sum_{l,r} \left(u_{l,t} \frac{\lambda_{l,r,t}}{\mu_{m \to l}} \right)^{\zeta}}. \tag{C.1}$$

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

$$\hat{\pi}_{m\to n,t} = \zeta \sum_{s\in S} \left\{ (1 - \pi^*_{m\to n})\pi^*_{s|n}\hat{\lambda}_{n,s,t} - \sum_{l\neq n} \pi^*_{m\to(l,s)}\hat{\lambda}_{l,s,t} \right\}, \tag{C.2}$$

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

$$\hat{\pi}_{n,t} = \sum_{m\in N} \frac{\pi^*_{m\to n}\pi^*_m}{\pi^*_n}\hat{\pi}_{m\to n,t}, \tag{C.3}$$

where $\hat{\pi}_{m\to n,t}$ is given by Equation (C.2). 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 (C.2), combined with Equation (C.3), 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.

D Details on estimating migration and diffusion frictions

D.1 Gravity equation for migration flows

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

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

To derive this equation, take the sum across all sectors s of the right-hand-side of Equation (C.1) 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 parameter $\zeta\bar{\mu}$. We restrict the sample to individuals in the IPUMS between the age of 20 and 60, 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 D.4: **Gravity equation for migration flows**

	Dependent var.: Migration probability
Distance in km	-0.00299***
	(0.00003)
Origin location FE	yes
Destination location FE	yes
# Obs.	228,484
Estimation	PPML

Notes: PPML estimates. The sample includes all pairs of 478 commuting zones in our main sample. The dependent variable is the migration probability between each pair of commuting zones. Standard errors in parenthesis. ***$p < 0.01$.

We estimate Equation (26) via Poisson Pseudo Maximum Likelihood (PPML), which is useful to accommodate zero values (there are 228 commuting zone pairs with migration probability equal

to zero). Results are reported in Table D.4. The estimate of the composite parameter, $\zeta\bar{\mu}$, is 0.003, implying that increasing distance by 100 kilometers decreases the migration probability by roughly 35% which, at the average migration probability, corresponds to 0.07 percentage points.

D.2 Gravity equation for knowledge flows

Consider now the gravity equation for knowledge flows (33) introduced and derived in Section 4.5:

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

To estimate the composite parameters $\theta\delta^G$ 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 $\eta^t_{(m,r)\to(n,s)}$ as the share of citations given by patents in (n,s) and directed to patents in (m,r).

Table D.5: **Gravity equation for knowledge flows**

	Dependent var.: Share of citations
Origin CZ $\neq$ Destination CZ	-7.6194***
	(0.0044)
Origin location-sector FE	yes
Destination location-sector FE	yes
Origin-Destination sector FE	yes
# Obs.	27,295,328
Estimation	PPML

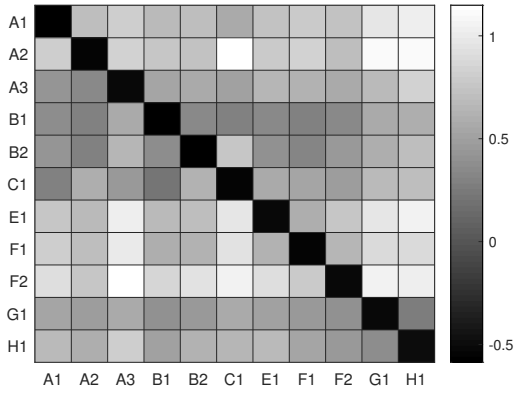
Notes: PPML estimates. The sample includes patents filed since 1980 and issued since 2000. Observations are all the combinations of pairs of location-sectors. The dependent variable is the share of citations given by each destination location-sector to each origin location-sector, where each citing patent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

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 local knowledge flows are more likely to be tacit and less likely to be captured by citations, we add to each patent's list of references an artificial citation to a local patent whose technology classes are for one half identical to the citing patent and for the remaining half identical to the local distribution of patents across technology classes in the period 1980-1999. This also

guarantees that all patents, including the ones with no backward citations, are included in the estimation.

We estimate this relationship by PPML and report results in Table D.5. The estimated coefficient combined with the value of θ implies $\delta^G = 3.63$. 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 D.13.

Figure D.13: **Knowledge transmission costs across sectors**



Notes: PPML estimates of $\delta^K_{r \to s}$, from regression of Appendix Table D.5. 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: 27,295,328. Class-groups are described in Appendix Table A.1.

E Robustness

In this section, we explore robustness of the model's main results to two choices of parameters explained in Section 4. First, we re-calibrate the model by setting ζ (the elasticity of migration with respect to average income) to the baseline value in Allen and Donaldson (2020), i.e., $\zeta = 4$. Second, we re-calibrate the model by setting ξ (the rate of knowledge obsolescence introduced in Section 4.4.1) to $\xi = 0.19$, which is consistent with the rate of knowledge obsolescence assumed by Anzoategui et al. (2019).[44]

Table E.6 reports the calibrated parameters under these two alternative assumptions. The only parameter that is significantly affected by these choices is the estimate of ω. In particular, lower migration elasticity requires weaker congestion forces to rationalize a zero correlation between change in exogenous amenities and exposure to technological waves. All the other parameters are only marginally affected.

Table E.6: **Parameter values and targets: robustness**

Parameter	Value			Target	Model	Data
	Main	$\zeta = 4$	$\xi = 0.19$			
σ	0.215	0.218	0.216	s.d. log-income (across CZs), 1990	0.191	0.191
θ	2.10	2.10	2.10	s.d. log-income (overall), 1990	0.67	0.67
A^o	2.11	2.11	2.11	Experience premium, 1990	1.42	1.42
ω	-0.092	-0.026	-0.099	Condition (31)	0.0	0.0
ζ	6.0	4.0	6.0	Assigned (Allen and Donaldson, 2020)		
$\bar{\mu}$	0.0005	0.0005	0.0005	Gravity equation (Table D.4)		
δ^G	3.63	3.63	3.63	Gravity equation (Table D.5)		
$\delta^K_{r\to s}$	Figure D.13			Gravity equation (Table D.5)		

Notes: Standard deviation of log-income for the overall population is taken from Krueger and Perri (2006). Experience premium is taken from Heathcote et al. (2010). Standard deviation of log-income across CZs are author's calculations from the NHGIS.

Table E.7 reports the main results for the two alternative calibrations, analogously as in Table 3 in the main paper. The main results are not substantially affected by these parameter choices. Columns 2-4 show that our mechanism explains 41.9% of the empirical relationship under $\zeta = 4$ and 52.3% under $\xi = 0.19$ (compared to 57.1% in the baseline calibration). When decomposing the effect (columns 5-7), we find that frictions to knowledge diffusion across technological fields explain 55.4% and 53.7% under $\zeta = 4$ and $\xi = 0.19$, respectively (compared to 64.0% in the baseline calibration). Overall, we find that our main results are robust to different parameter choices for ζ and ξ.

[44] Anzoategui et al. (2019) assume a quarterly rate of obsolescence of 2%, as the average of the estimates from Caballero and Jaffe (1993) and Bosworth (1978).

Table E.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		
		Main	$\zeta = 4$	$\xi = 0.19$	Main	$\zeta = 4$	$\xi = 0.19$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposure to tech. waves ($Exp_{n,t}$)	0.327***	0.140	0.190**	0.156*	0.207**	0.251***	0.235***
	(0.085)	(0.085)	(0.085)	(0.084)	(0.084)	(0.085)	(0.083)
Difference from empirical coefficient	-	0.188	0.137	0.171	0.121	0.076	0.092
Share explained by tech waves ($\alpha_{s,t}$)		57.1%	41.9%	52.3%	36.5%	23.2%	28.1%
Decomposition:							
- Share explained by tech. frictions					64.0%	55.4%	53.7%
- Share explained by geo. frictions					36.0%	44.6%	46.3%
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
# Obs.	2,818	2,818	2,818	2,818	2,818	2,818	2,818
R^2	0.52	0.50	0.50	0.50	0.51	0.51	0.50

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 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. $^{**}p < 0.05$, $^{***}p < 0.01$.

F Deriving the specialization measure

To rationalize the measure of specialization in Equation (35), 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 we derive the following theoretical result, that links the volatility of local population growth to the local degree of specialization:

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

$$Var(\hat{\pi}_{n,t}) \propto \sum_{s \in S} \left(\pi^*_{s|n} - \pi^*_{\cdot,s}\right)^2. \tag{F.4}$$

Proof. Under Assumption A2.1, 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:

$$\hat{\pi}_{n,t} \overset{A2.1}{\propto} (1 - \pi^*_n) \sum_{s \in S} \left(\pi^*_{s|n} - \pi^*_{s|-n}\right) \hat{\alpha}_{s,t}. \tag{F.5}$$

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

$$Var(\hat{\pi}_{n,t}) \overset{A2.1,A5}{\propto} (1 - \pi^*_n)^2 \sum_{s \in S} \left(\pi^*_{s|n} - \pi^*_{s|-n}\right)^2.$$

Under Assumption A2.2, all cities are negligible in size compared to the overall economy. Hence, the variance of the percentage change in the population share of location n is equal to the measure of specialization in Equation (35):

$$Var(\hat{\pi}_{n,t}) \overset{A2,A5}{\propto} \sum_{s \in S} \left(\pi^*_{s|n} - \pi^*_{\cdot,s}\right)^2 \equiv Spec_n.$$

□

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 (2020): "Persistence and path dependence in the spatial economy," Tech. rep., National Bureau of Economic Research.

Anzoategui, D., D. Comin, M. Gertler, and J. Martinez (2019): "Endogenous Technology Adoption and R&D as Sources of Business Cycle Persistence," *American Economic Journal: Macroeconomics*, 11, 67–110.

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

Bosworth, D. L. (1978): "The rate of obsolescence of technical knowledge–A note," *The Journal of Industrial Economics*, 273–279.

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.

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.

Krueger, D. and F. Perri (2006): "Does income inequality lead to consumption inequality? Evidence and theory," *The Review of Economic Studies*, 73, 163–193.

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.