Cities and Technological Waves*

Enrico Berkes† Ruben Gaetani‡ Martí Mestieri§

10th March 2021

Abstract

We develop a spatial equilibrium model of economic growth with idea diffusion to quantify the role of technological waves in determining the evolution of the U.S. economic geography from 1890 through 2010. Leveraging a comprehensive dataset of historical geolocated patents, we find that changes in the technological environment coupled with frictional idea diffusion explain more than half (58%) of the variation in U.S. city growth since 1890. The calibrated model reproduces the rise and fall of the Rust Belt and the emergence of modern knowledge hubs, and implies heterogeneous geographical effects of alternative future technological scenarios.

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

*We thank Mike Andrews, Paco Buera, Ben Jones, Joel Mokyr, Nicola Persico, Frédéric Robert-Nicoud, Bruce Weinberg, our discussants Klaus Desmet and Ed Glaeser, and seminar attendees at Northwestern, University of Toronto, Ohio State University, Conference of Swiss Economists Abroad, NBER Summer Institute (Urban Economics), VMACS Junior Conference, and 2021 AEA Meetings 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.

†Ohio State University, berkes.8@osu.edu.
‡University of Toronto, ruben.gaetani@utoronto.ca.
§Federal Reserve Bank of Chicago, Northwestern University and CEPR, marti.mestieri@northwestern.edu.

1 Introduction

The economic geography of countries is in perpetual evolution. In the United States, many cities and regions that have thrived in the past have progressively lost population and influence in favor of newly emerging areas. In recent decades, several cities in the Rust Belt, that had experienced extraordinary growth throughout most of the 20th century, have entered a prolonged phase of decline. At the same time, a handful of urban areas specialized in knowledge-intensive sectors, such as information technology and pharmaceuticals, have gained prominence, becoming increasingly attractive for workers and firms (Glaeser and Gottlieb, 2009; Moretti, 2012). The determinants of these rich dynamics in city growth are still a matter of debate and remain a central question in urban economics.

In this paper, we develop a quantitative framework to study the impact of *technological waves*, defined as long-term swings in the importance of sectors in the innovation space, on the evolution of the U.S. economic geography over the last 150 years. To measure technological waves and their effect on the growth trajectory of cities, we leverage a new dataset of historical U.S. patents geolocated at the city level spanning the period 1870 through 2010. The data reveal a robust positive relationship between a city's exposure to technological waves and its ability to attract population over the following decades. We propose that frictions in the diffusion of ideas across space and fields of knowledge prevent cities from optimally reallocating resources towards expanding technologies. Using the same data, we document persistently localized patterns of patent citations – both in the geographical and knowledge space – that support this argument and indicate that a city's ability to embrace new technological opportunities is constrained by the local availability of complementary ideas.

Motivated by these findings, we formalize the feedback between technological waves and the evolution of the economic geography in a quantitative spatial-equilibrium model with innovation and frictional knowledge diffusion. The model remains tractable and retains a recursive structure that allows us to calibrate it using our patents and population data in combination with a small set of transparent assumptions. We find that technological waves combined with frictional knowledge diffusion account for more than half (58%) of the variation in U.S. city growth since 1870. The calibrated model reproduces the rise and fall of the Rust Belt and the emergence of modern knowledge hubs in the last decades of the 20th century, and implies a quantitatively sizeable role for local diversification in making cities more resilient to changes in the technological landscape. To conclude, we investigate how alternative scenarios of future technological waves might transform the U.S. economic geography. The model predicts substantial differences in the geographical effects of scenarios such as the rise of autonomous vehicles, and the expansion of medical sciences or sustainable agriculture.

The model combines an economic geography setting with a theory of economic growth that emphasizes the role of recombination, imitation, and knowledge diffusion, as recently developed by Lucas and Moll (2014), Perla and Tonetti (2014), and Buera and Oberfield (2020) among others. In the model, newborn agents make migration and occupational decisions after forming expectations on their lifetime productivity in the location and sector of their choice. Productivity is determined by an imitate-or-innovate decision. Agents can either imitate an idea drawn from the local knowledge distribution, or innovate by improving upon an idea drawn from the distribution of any other location and sector in the economy. The applicability of an idea is affected by frictions reflecting both geographical and technological distance. These frictions imply that knowledge drawn *within* any location-sector can be converted into new inventions more effectively than knowledge drawn from other locations and sectors. For this reason, a city's stock of knowledge determines not only current productivity but also future innovation possibilities, making the local growth trajectory sensitive to technological wave shocks. To focus on this novel interplay between economic geography and idea diffusion, the model purposefully abstracts from other drivers of city growth such as endogenous residential amenities.

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 shocks, the model features a unique balanced growth path (BGP). The productivity distribution for each location-sector endogenously retains a Frećhet structure, and implies an intuitive equation for the law of motion of its scale parameter. This also allows us to characterize knowledge flows in closed form through a gravity representation that can be estimated using patent citation data.

Despite the relative parsimony of the model, the linkages across space and fields of knowledge that it generates imply non-trivial population dynamics. Before turning to the quantitative analysis, we study the mechanics of the model by log-linearizing the equilibrium conditions around the BGP. We use the log-linear dynamics to derive theoretical predictions on the relationship between technological waves, the evolution of local productivity, and city growth. First, the growth rate of the scale parameter of the knowledge distribution in each location-sector can be expressed as the sum of technological wave shocks weighted by the local reliance on the idea's sector of origin. This implies that cities specialized in expanding (declining) sectors will experience higher (lower) local productivity growth. Second, if knowledge flows across sectors are of second-order importance relative to flows within sectors, a measure of local exposure to technological waves relative to the overall economy – similar to a shift-share variable – is a sufficient statistic to predict local population growth. In particular, a city grows if and only if technological wave shocks weighted by the incidence of each sector in the city are larger than the corresponding weighted average for the rest of the economy.

We then turn to the quantitative assessment of the role of technological waves – and their interaction with frictional knowledge diffusion – in explaining the evolution of the U.S. economic geography since 1890. We show that the model has a recursive structure that allows us to calibrate the parameters and to recover the unobserved disturbances – including the technological wave shocks – by imposing a small set of transparent assumptions.

The calibrated model is successful in capturing key features of the data that are not directly targeted, and suggests that the endogenous mechanism of knowledge creation and diffusion, interacted with the estimated technological waves, can account for 58% of the variation in population growth across U.S. cities between 1890 and 2010. The equation for the dynamics of the local stock of knowledge allows us to isolate a structural residual that captures all the factors affecting the evolution of local innovation that cannot be directly ascribed to the endogenous mechanism of knowledge creation and diffusion. These factors include a variety of forces that can be either exogenous (e.g., natural events) or endogenous (e.g., opening of new research facilities) with respect to the local exposure to technological waves. The framework does not require to make any assumptions on the nature of this structural residual, but allows us to discern cases in which residual factors amplify or dampen the direct effect of technological wave shocks.

We further show that the model is successful in accounting for two of the most prominent transformations of the U.S. economic geography of the last century: the rise of manufacturing-intensive cities in the early decades of the 20th century, their later decline, and the subsequent emergence of knowledge hubs specialized in information technology. We find that the mechanism of endogenous knowledge creation and diffusion can explain a significant portion of the growth (and subsequent decline) of the major centers of heavy manufacturing, with residual forces amplifying the oscillations in their growth trajectory. This experience was mirrored in recent decades by some of the most rapidly expanding innovation centers in the United States.

The mechanism of knowledge creation and diffusion also implies that the degree of local diversification has a central role in determining a city's resilience to technological waves. Simulations of counterfactual paths of technological waves reveal that more diversified cities experience significantly less volatile growth trajectories. There are two factors behind this relationship. First, even in the absence of knowledge flows across fields, more diversified cities have a lower chance of large exposure to technological wave shocks (either on the positive or negative side) since negative shocks to some sectors are likely to be compensated by positive shocks to other sectors. Second, accounting for knowledge flows across fields, more diversified cities have a broader availability of ideas to draw from, implying that (positive or negative) shocks to individual sectors have a lower impact on the evolution local productivity.

Finally, we use the quantitative model to predict how the U.S. economic geography will evolve in the coming decades under different scenarios for the evolution of the technological landscape. In particular, we study which cities benefit – and which do not – compared to the technological status quo, in the following scenarios: (1) a rise 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 find that cities in the Rust Belt benefit from the first scenario, at the expense of cities in the North-East and the Pacific. The second scenario penalizes knowledge hubs specialized in IT-related innovation, favoring more diversified areas such as Boston and the cities in California outside the Silicon Valley. The third scenario prompts a reallocation of economic activity towards the agricultural areas in the Central states.

Related Literature This paper contributes to multiple strands of the literature. First, the theory is based on idea flows at the location-sector level, with technological and geographical frictions in knowledge diffusion playing a key role in explaining city dynamics. While a rich body of literature has documented the strength and geographical span of localized knowledge spillovers (among others, Jaffe et al., 1993; Audretsch and Feldman, 1996; Greenstone et al., 2010) there has been no attempt to perform a quantitative assessment of the importance of these externalities for understanding long-run city dynamics. One of the main obstacles for providing such an assessment is the complexity of modeling idea diffusion in a spatial setting. In recent years, two flourishing bodies of literature have provided major methodological advances in this direction. First, a number of papers have developed tractable endogenous growth models that emphasize recombination, imitation, and knowledge diffusion as major drivers of aggregate productivity growth (e.g., Perla and Tonetti, 2014; Lucas and Moll, 2014; Buera and Oberfield, 2020). Second, a rich body of work on quantitative spatial economics has developed tools for studying the distribution of economic activity in space, 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 strands of the literature and develops a dynamic, multi-sector, endogenous growth model in a spatial economy that is highly tractable and can be quantitatively disciplined using data on population and patents over a long time period. While a number of papers have used detailed data on patenting to study innovation and knowledge flows in firm and industry dynamics (e.g., Akcigit and Kerr, 2018; Cai and Li, 2019), or developed static models that emphasize localized knowledge spillovers as the main determinant of the economic geography (e.g., Davis and Dingel, 2019), this paper is, to the best of our knowledge, the first attempt at quantitatively assessing the importance of frictions in knowledge diffusion for city dynamics.

[1] Comprehensive reviews of these bodies of literature are provided by Buera and Lucas Jr (2018) for models of endogenous growth with idea flows, and by Redding and Rossi-Hansberg (2017) for quantitative spatial equilibrium models.

An extensive literature has investigated the forces governing the long-run evolution of the economic geography, specifically in its propensity to display path dependence and occasional reversal of fortune (e.g., Brezis and Krugman, 1997; Davis and Weinstein, 2002; Bleakley and Lin, 2012; Kline and Moretti, 2014), as well as in its responsiveness to aggregate shocks such as rising sea-level (e.g., Desmet et al., 2018a), and regional or sectoral shocks (e.g., Caliendo et al., 2018; Hornbeck and Moretti, 2018). The working hypothesis in this paper is that aggregate changes in the technological landscape, combined with frictional knowledge transmission, have a first-order impact on the geographical distribution of economic activity. The framework can account simultaneously for path dependence and reversal of fortune in city dynamics. While the focus on innovation and idea diffusion is new to this literature, there is a rich body of work that has analyzed the historical dynamics of the U.S. geography, both from an empirical perspective (e.g., Bostic et al., 1997; Simon and Nardinelli, 2002; Desmet and Rappaport, 2017) and from a structural and quantitative viewpoint (e.g., Desmet and Rossi-Hansberg, 2014; Nagy, 2017; Allen and Donaldson, 2018; Eckert and Peters, 2019).

This paper also contributes to the long-standing debate between the returns to local specialization (Marshall, 1890) and urban diversity (Jacobs, 1969), and their effect on city growth. Notable contributions in this literature include Glaeser et al. (1992), whose empirical assessment finds evidence supporting Jane Jacob's view of urban variety 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 suggests and quantifies a new channel through which urban diversification affects long-run city growth, namely, the responsiveness of a city to changes in the surrounding innovation landscape. It also implies a tradeoff between larger growth opportunities during favorable waves and more severe downturns during adverse ones.[3] In this sense, the model provides a new lens for interpreting the effect of local policies directed at increasing local diversity.

The remainder of the paper is organized as follows: Section 2 introduces the data and presents historical trends and 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 calibration and Section 5 presents the quantitative results. Section 6 discusses avenues for further research and concludes.

[2]A comprehensive overview of the patterns of specialization across U.S. locations is provided by Holmes and Stevens (2004).

[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

Technological change is a slow-moving secular process. To study how the rise and fall of technologies determines the growth and decline of cities, we therefore need to consider a time period long enough to capture multiple episodes of technological replacement. In this paper, we exploit a recently assembled dataset of historical patents spanning a period of almost two centuries to measure technological waves and the position of cities in the innovation space. We approximate cities using a full partition of the U.S., namely the 1990 commuting zones (CZs), that we keep fixed throughout the analysis.[4]

2.1 Data sources

To measure innovative activities at the city level, we collect patents data from the Comprehensive Universe of U.S. Patents (CUSP). The CUSP contains information about patents filed (and subsequently issued) by the U.S. Patent and Trademark Office (USPTO) between 1836 and 2010, with an estimated coverage above 90% in each year.[5] Particularly, the CUSP provides information about the technology classes, name and location of each inventor (and assignee) listed on a patent, as well as their filing and issue dates. It assigns patents to the city of residence of each inventor and does not rely on the county reported in the patent's text. This allows us to build geographically consistent measures of innovation over almost two centuries.

To construct consistent population measures for 1990 commuting zones, 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.[6] 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.[7] Third, we assign the resulting town popula-

[4]Although commuting flows have certainly changed over time, assuming a stable geography allows us to abstract from annexations and redefinition of town borders that have been pervasive phenomena throughout the 19th and 20th century.

[5]Berkes (2018) provides details about the data collection procedure, as well as summary statistics and stylized facts related to the underlying data. Andrews (2019), in a comparison of historical patents data, describes it as "currently the gold standard, in terms of the patent- and inventor-level information included in the published datasets". Some slices of the data have already been used in Berkes and Nencka (2020) who study the effect of Carnegie libraries on the local patenting activity, Clemens and Rogers (2020) who study how procurement policies affect the characteristics of medical innovation, and Babina et al. (2020) who study the effect of the Great Depression on innovative activities in the U.S.

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

[7]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.

tion to 1990 commuting zones.[8]

Following the same approach, we construct consistent measures of human capital that combine available information on local literacy and education over time. To make this measure comparable across decades, we rank cities in terms of the relevant measure for that decade and use the resulting ranking for the analysis. To the best of our knowledge, this paper is the first to construct consistent local measures of human capital over a time span of over a century.

In both the empirical analysis and the model, we restrict the sample to the subset of commuting zones in the contiguous United States that accounted for at least 0.02% of the total population for each decade since 1890. This delivers a sample of 373 commuting zones, that jointly account for roughly 87% of the U.S. population in 2010.[9]

Throughout the empirical and structural analysis, observations correspond to 20-year periods between 1890 and 2010. For each period, we assign demographic characteristics from the corresponding decennial census (ACS for 2010), and compute patent counts by sector by adding patents filed in the two decades around the focal year (for example, patents in the 1990 observation correspond to the total patent count between 1980 and 1999). Sectors are defined as the technological class-groups obtained by grouping 3-digit IPC classes into 11 class-groups, as detailed in Appendix Table A.1.

2.2 Historical trends

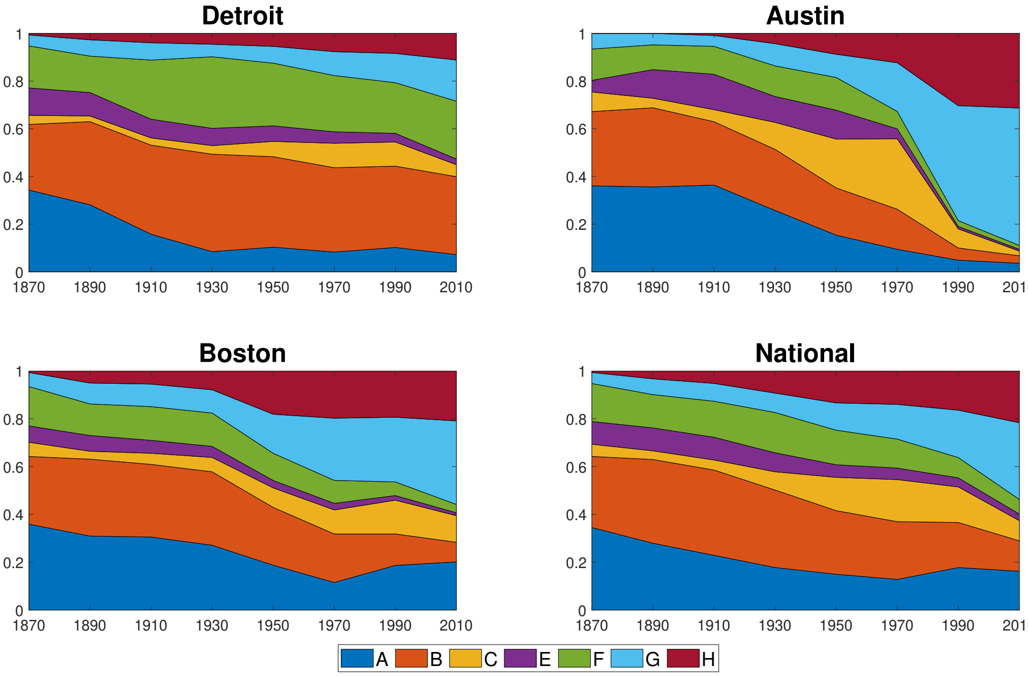
The last 150 years have witnessed major shifts in the technological landscape. The bottom-right panel of Figure 1 shows how the distribution of the national patenting output across the 7 main IPC classes has evolved since 1870.[10] The share of patents in class A ("Human necessities") – that includes innovation related to both agriculture and medical sciences – declined in the first part of the century, as agriculture lost its centrality to classes complementary to the heavy manufacturing industry, such as B ("Performing Operations; Transporting") and F ("Mechanical Engineering"). Class A rebounded in recent decades as innovation in medicine gained prominence. In the second part of the century, classes G ("Physics") and H ("Electricity") became more central in the national shares, making up more than 50% of the overall

[8] 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.

[9] As a reference point, this rule requires that cities had a population of at least 10,711 people in 1890 and 60,387 people in 2010.

[10] The full description of each class is available at https://www.wipo.int/classifications/ipc/en/.

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



Notes: Composition of patenting output across the 7 main IPC classes in Appendix Table A.1. Patent count for year t is constructed as the sum of patents filed between $t-10$ and $t+9$. A: Human necessities. B: Performing operations; Transporting. C: Chemistry; Metallurgy. E: Fixed Constructions. F: Mechanical Engineering; Lighting; Heating; Weapons; Blasting. G: Physics. H: Electricity.

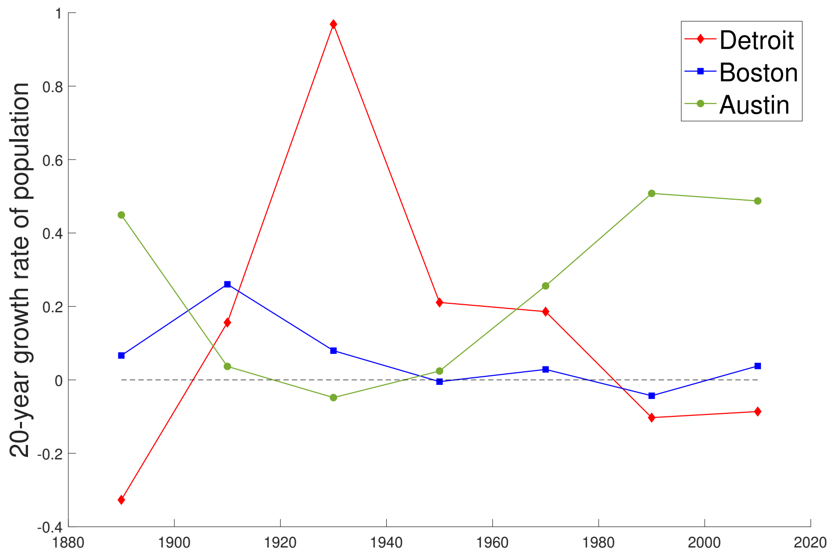
innovation output in 2010.[11]

The composition of patenting not only changes significantly over time, but also varies considerably across cities at any point in time. The top panels of Figure 1 depict two of the archetypal examples of this heterogeneity. Detroit (top-left) has been specialized in the production of patents of class B and F since the early 1900s. In 1930, these two classes made up about 70% of its patenting portfolio. This pattern has remained broadly unchanged throughout the century, with a slight shift towards patents of classes G and H since the 1990s. Austin (top-right) exhibits fairly diversified innovation activities until the 1970s, when the share of patents of classes G and H started expanding, reaching 90% of the portfolio by 2010. By contrast, Boston (bottom-left) displays a diversified patenting output that, throughout the decades, has closely tracked the national trends.

In this paper, we argue that the heterogeneity in the composition of local patenting, combined with frictions to knowledge diffusion, makes cities unevenly positioned to take advantage of new innovation opportunities. This makes cities' trajectories sensitive to changes in the

[11] Classes G and H include the bulk of innovation related to computers, electronics, and information and communication technology.

Figure 2: **City dynamics**



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

technological landscape, and contributes to explaining the irregular historical dynamics of U.S. urban and regional growth. The experiences of Detroit, Austin, and Boston since the late 1800s exemplify this point. Figure 2 shows the 20-year population growth of those three commuting zones since 1890, after controlling for Census Division-time fixed effects. Detroit displays the most striking growth rates in the decades after the advent of the automobile industry around 1910, followed by a long-lasting decline that resulted in a steady loss of population since the 1980s. The commuting zone of Austin experienced a specular trajectory. The city lost population in the first half of the 20th century, as the Texas Oil Boom favored areas of the state that were rich of oil, making Austin slip from the 4th to the 10th place among Texas's largest cities.[12] However, in recent decades Austin has emerged as one of the leading innovation hubs in the country, leveraging its richness of science-based firms and a large college-educated population. Finally, the commuting zone of Boston had a still different experience: Throughout the last century, it has retained a considerably less volatile path, characterized by moderate population growth interrupted by occasional periods of modest population decline. The consistent diversification of Boston's patenting output could have made the city less sensitive to technological waves, explaining the stability of its growth path.[13]

[12] https://tshaonline.org/handbook/online/articles/hda03

[13] 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.

2.3 Technological waves and the growth and decline of cities

Taken together, Figure 1 and 2 suggest that changes in the importance of technological fields, here captured by shifts in the national shares, might differentially affect the growth trajectory of cities based on their patenting portfolios. We now provide a systematic assessment of these patterns.

In the spirit of Bartik (1991), we define commuting zone n's exposure to the technological wave in decade t as the sum across all class-groups $s \in S$ of the growth rate in the national share of each class-group, weighted by the corresponding local share of the same class-group in the previous period:

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

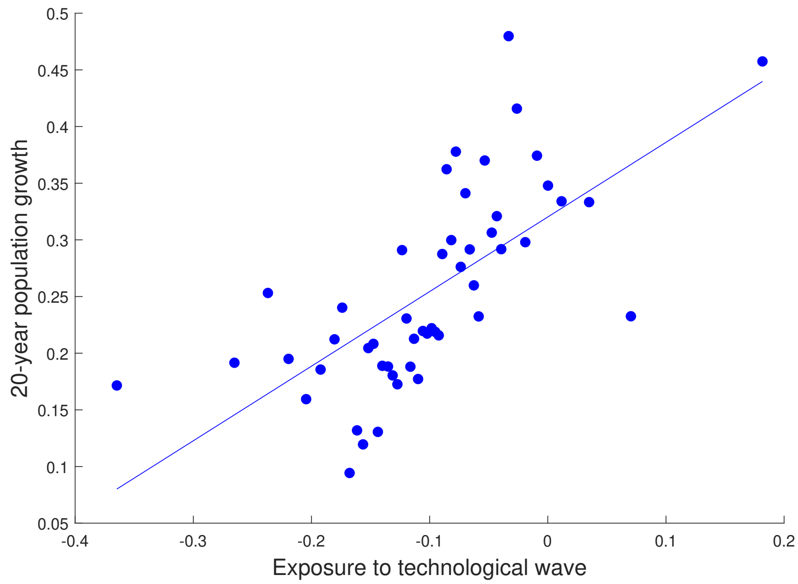
In Equation (1), $Share_{n,s,t-1}$ is the share of patents filed in commuting zone n of class-group s at time $t-1$, and $g_{s,t}$ is the growth rate in the national share of patents of class-group s between $t-1$ and t. Positive (negative) values of $Exp_{n,t}$ imply that the initial patenting output of n is concentrated in class-groups whose national share is expanding (shrinking).

Figure 3 shows a bin-scatter plot of the relationship between the measure of exposure to the technological wave, $Exp_{n,t}$, and the 20-year growth rate of local population between 1890 and 2010. Both measures are residualized with respect to lagged log-population and Census Division-time fixed effects, to account for size and convergence effects and for the differential growth rates of commuting zones across space explained by factors such as the Westward expansion or the Great Northward Migration. The scatter plot reveals a strong positive correlation, implying that, over the period considered, cities with a more favorable exposure to the technological wave have experienced higher population growth than cities in the same Census Division.

Table 1 reports the regression results. The estimate in Column 2 (the baseline specification with Census Division-time fixed effects) implies that an increase in the measure of exposure of one residual standard deviation is associated with an increase of 14.7% of a residual standard deviation in population growth. In Column 3, we further control for a historically-consistent measure of local density of human capital.[14] This indicator is correlated with population growth, but has a negligible effect on the estimated coefficient of the exposure measure. In Appendix Table A.3, we show that results are consistent when splitting the sample into early (1890-1930) and late (1950-2010) periods.

[14] This measure corresponds to a within-decade ranking of CZs along a summary index that includes several measures of human capital. The specific indicators we use change over time depending on the availability of information in the historical Census. In the early decades, the measure focuses on indicators of literacy and schooling, while in later decades it emphasizes the local density of workers with high educational attainment.

Figure 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, 1890-2010. The bin-scatter plot is residualized with respect to Census Division-time fixed effects and lagged log-population.

2.4 Frictions to knowledge diffusion across locations and fields of knowledge

In this paper, we propose that the robust correlation in Table 1 is partly due to the existence of geographical and technological frictions to the diffusion of ideas, that prevent cities from optimally reallocating resources to take advantage of technological waves. As a result, cities whose innovation portfolio is heavily skewed towards expanding fields are better positioned to embrace new innovation opportunities and will become more attractive for workers and firms.

The fact that knowledge diffusion is highly localized – both geographically and in the knowledge space – has been widely 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 patterns of patent citations (Murata et al., 2014, Kerr and Kominers, 2015). In Figure 4 we show that the localization of patent citations is visible at our level of geographical and technological disaggregation, and over the long time span covered by our patents data.

The heatmap in the left panel of Figure 4 displays, for each technology class-group on the vertical axis, the probability that a citations from that class-group fall under each of the class-groups on the horizontal axis. Since patent citations are not consistently available in the earlier

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

	Growth rate of population		
	(1)	(2)	(3)
Exposure to tech. wave	.588***	.521***	.468***
	(.143)	(.143)	(.143)
Lagged log-population	-.078*	-.057	-.070
	(.044)	(.045)	(.046)
Human capital (ranking)			.187***
			(.067)
Fixed effects	T	CD×T	CD×T
# Obs.	2,611	2,611	2,611
R^2	0.24	0.39	0.40

Notes: CZ level regression, 1890-2010. 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. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

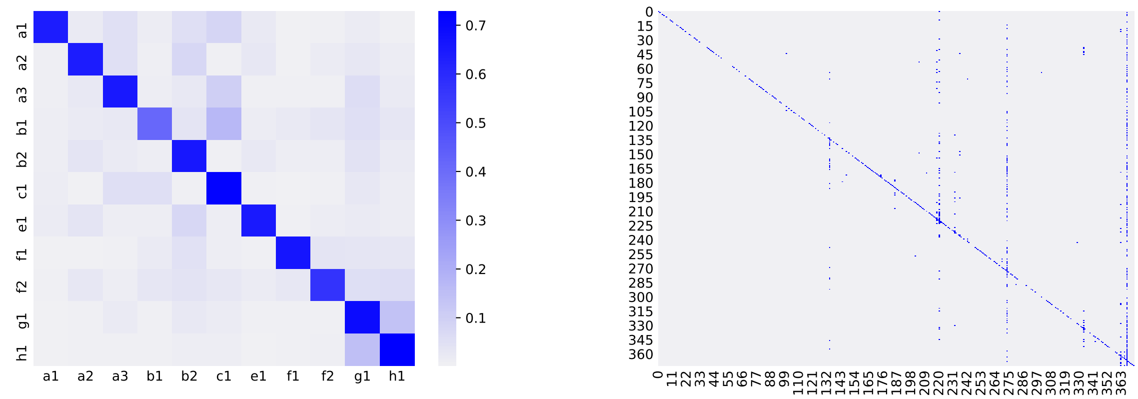
decades, we restrict the sample to all the patents filed since 1950.[15] The right panel displays the corresponding citation probabilities for each pair of commuting zones.

Both heatmaps show that citations are strongly concentrated along the diagonal, suggesting a high degree of technological and geographical localization in the diffusion of ideas. The heatmaps also reveal that some of the class-groups and commuting zones are more likely to be cited than others, reflecting a combination of higher size (number of grants) and higher generality. In the quantitative model, we derive a gravity equation for knowledge flows that separates the effect of size on bilateral citation probabilities from that of localization and generality. We use the estimated gravity equation to discipline the parameters controlling the frictions to idea diffusion.

Appendix Figure B.1 shows that the patterns of localization remain strongly visible when splitting the data in an early (1950-1979) and a late (1980 onwards) samples. This suggests that, while improvements in communication technologies might have made the diffusion of ideas across space more effective, they did not fundamentally alter the strongly localized nature of knowledge flows.

[15] A separate section containing referenced patents was formally introduced only in 1947.

Figure 4: **Patent citations across fields of knowledge (left) and locations (right)**



Notes: Left panel: Probability that patents from the class-group on the vertical axis cite patents from the class-group on the horizontal axis. Right panel: Probability that patents from the commuting zone on the vertical axis cite patents from the commuting zone on the horizontal axis. Only patents filed after 1950 are included in the sample.

3 Model

In this section, we develop a structural model that embeds endogenous growth through innovation and frictional idea diffusion into a quantitative spatial equilibrium framework. The theory formalizes the feedback between changes in the innovation landscape and the evolution of the economic geography over time, and rationalizes the reduced-form relationship between population growth and exposure to technological waves. The model has a recursive structure that allows us to describe a natural mapping between the model's objects and the data (including a gravity equation for knowledge flows). We use the calibrated model to quantitatively assess the importance of technological waves and frictional knowledge diffusion in explaining the historical dynamics of the U.S. economic geography, and to speculate on its evolution in future decades under plausible scenarios of technological trends.

3.1 Environment

We consider an economy comprising a finite set N of locations and a finite set S of sectors. In what follows, we refer to N and S as both the sets of locations and sectors, and their cardinality. Time is discrete and indexed by t. At each point in time, the economy is populated by a mass L_t of individuals.

3.1.1 Preferences, endowments, and demographics

In each period, a new generation of individuals is born in the location of their parents and makes migration and occupational decisions. Individuals live for one period and, at the end of the period, have f_t children. There are no moving costs.

Migration and occupational choices are made to maximize expected utility, subject to idiosyncratic utility draws that affect the individual desirability of each location-sector pair. Specifically, at the beginning of the period, each individual i receives a full set of stochastic utility draws, one for each location-sector in the economy:

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

Each value $x_{n,s,i}$ is a random draw from a Fréchet distribution with shape parameter $\zeta > 1$. Individuals then choose the location-sector pair (n, s) that provides them with the highest expected utility, given by:

$$U_{n,s,t}(\mathbf{x}_i) = u_n\, x_{n,s,i}\, c_{n,s,i,t}, \tag{2}$$

where u_n is the level of time-invariant 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.

Since we calibrate the time period to be 20 years, the assumptions on the absence of moving costs and the demographic structure should be interpreted with this time horizon in mind. The model can be easily extended to account for overlapping generations and costly moving decisions.[16]

3.1.2 Production and innovation technologies

Each agent i is endowed with one unit of labor that she supplies inelastically with productivity q_i. Total output in the economy is given by a linear aggregator over individual productivity across all locations and sectors:

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

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

Individual productivity is determined endogenously by a process of knowledge diffusion that subsumes a choice on whether to imitate or innovate. At the beginning of each period every

[16] A previous version of this model, that allows for costly migration and overlapping generations, is available upon request.

agent i in the new generation receives a full set of idiosyncratic, independently distributed draws:

$$\mathbf{z}_{n,s,i} = \left\{ z^l_{n,s,i}, \left\{ z^x_{m,r,i} \right\}_{m,r \in N \times S} \right\}. \tag{3}$$

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 knowledge that individual i learns from their teacher, mentor, or manager, and can be imitated and adopted directly in production.[17] If the agent chooses to adopt this idea in production, their lifetime productivity is

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

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 $\{F_{m,r,t}(q)\}_{m,r \in N \times S}$. Note that this full set of draws includes local ones (i.e., $m = n$ and $r = s$). These draws can be interpreted as knowledge that the agent acquires by various channels of transmission, such as books, radio, television, internet, or via casual interactions with local or non-local individuals. Although these ideas cannot be imitated and adopted directly in production, they can be used 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

$$q_{n,s,i,t} = \frac{\epsilon_{n,s,t}\, \alpha_{r,t}\, z^x_{m,r,i}}{d_{(m,r)\to(n,s)}}. \tag{4}$$

In Equation (4), the term $\alpha_{r,t}$ represents the centrality of sector r in the innovation landscape. The higher the value of $\alpha_{r,t}$, the more effectively can knowledge in sector r be developed into innovation for any sector. We refer to changes in $\alpha_{r,t}$ as "technological wave" shocks. The term $d_{(m,r)\to(n,s)}$ captures the geographical and technological frictions that discount the effectiveness of knowledge transmission between the idea origin (m, r) and the idea destination (n, s). The term $\epsilon_{n,s,t}$ is a structural residual that captures the current effectiveness of innovation in (n, s) and is common to all innovators in the location-sector pair. It accounts for all the residual factors that affect the productivity of the local sector but are not otherwise included in (4), such as the opening of production facilities, universities, and research centers.

There is no market to smooth consumption across generations. Thus, agents live hand-to-

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

mouth, with consumption of final good given by own production:

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

3.2 Equilibrium

3.2.1 Diffusion of knowledge

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

$$q_{n,s,i,t} = \max \left\{ z^l_{n,s,i},\ \max \left\{ \frac{\epsilon_{n,s,t}\, \alpha_{r,t}\, z^x_{m,r,i}}{d_{(m,r)\to(n,s)}} \right\}_{m,r \in N\times S} \right\} \tag{5}$$

Equation (5) shows how this process can be divided in two steps. First, the agent chooses the best innovative idea available to her. Then she compares this idea with her imitation draw, and picks the one that yields higher productivity for her.

The following assumption will play an important role in keeping the theory tractable:

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

$$F_{n,s,0}(q) = e^{-\lambda_{n,s,0} q^{-\theta}}. \tag{6}$$

A multivariate Fréchet distribution with common shape parameter is max-stable. This implies that, under Assumption 1, the resulting distribution over the max of Fréchet draws is also Fréchet with the same shape parameter.[18] Combining (5) with (6), we find that individual productivity at any time $t \geq 0$ is distributed Fréchet with shape parameter $\theta > 1$ and with scale parameter evolving according to the following law of motion:

$$\lambda_{n,s,t} = \underbrace{\lambda_{n,s,t-1}}_{\text{Imitation}} + \underbrace{\sum_{m\in N} \sum_{r \in S} \lambda_{m,r,t-1} \left(\frac{\epsilon_{n,s,t}\, \alpha_{r,t}}{d_{(m,r)\to(n,s)}} \right)^{\theta}}_{\text{Innovation}}. \tag{7}$$

Equation (7) summarizes the growth dynamics implied by the model. The scale parameter of the new generation in location-sector (n, s) is equal to the scale parameter of the previous generation augmented by a second term which captures inventive activities. This second term in Equation (7) is composed by the sum of scale parameters across all location-sectors weighted by their applicability to location-sector (n, s). This applicability term includes the importance of

[18]The same degree of tractability can be achieved without assuming independence, as in Lind and Ramondo (2019).

each field of knowledge ($\alpha_{r,t}$) and the local effectiveness of innovation ($\epsilon_{n,s,t}$), and is discounted by technological and physical distance between location-sector pairs ($d_{(m,r)\to(n,s)}$).

Equation (7) also implies that, conditional on innovating, the probability that an inventor in location-sector (n,s) builds upon an idea from any location-sector (m,r) at time t can be expressed as follows:

$$\eta^t_{(m,r)\to(n,s)} = \frac{\lambda_{m,r,t-1}\left(\frac{\alpha_{r,t}}{d_{(m,r)\to(n,s)}}\right)^{\theta}}{\sum_{l,p}\lambda_{l,p,t-1}\left(\frac{\alpha_{p,t}}{d_{(l,p)\to(n,s)}}\right)^{\theta}}. \tag{8}$$

3.2.2 Migration and occupational choice

At the beginning period t, agents in the new generation observe sectoral and local shocks ($\{\alpha_{r,t}\}_{r\in S}$ and $\{\epsilon_{n,s,t}\}_{n,s\in N\times S}$) but do not know their idiosyncratic idea draws, so they have to form expectations about productivity before making their migration and occupational decisions. Agent i moving to location-sector pair (n,s) has expected utility equal to

$$\mathbb{E}\left[U_{n,s,t}(\mathbf{x}_i)\right] = u_n\, x_{n,s,i}\, \mathbb{E}\left[q_{n,s,t}\right]. \tag{9}$$

In equilibrium, $q_{n,s,t}$ is distributed Fréchet with shape parameter θ and scale parameter $\lambda_{n,s,t}$, which can be inferred at time t via the law of motion (7), so that

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

where $\Gamma(\cdot)$ denotes the gamma function. This implies that the probability that any newborn individual selects location-sector (n,s) is

$$\pi_{n,s,t} = \frac{\left(u_n\,\lambda_{n,s,t}^{\frac{1}{\theta}}\right)^{\zeta}}{\sum_{m,r}\left(u_m\,\lambda_{m,r,t}^{\frac{1}{\theta}}\right)^{\zeta}}. \tag{11}$$

Thus, the mass of agents in location-sector (n,s) at time t is equal to

$$L_{n,s,t} \equiv \pi_{n,s,t}\, L_{t-1}\, f_t. \tag{12}$$

For notational convenience, we define $\pi_{n,t} \equiv \sum_{s\in S}\pi_{n,s,t}$ and $L_{n,t} \equiv \sum_{s\in S} L_{n,s,t}$ as, respectively, the share and mass of individuals living in location n.

3.2.3 Equilibrium Definition

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

Definition 1. *For a given set of initial conditions*

$$L_0,\ \{u_n\}_{n\in N},\ \{\lambda_{n,s,0}\}_{n,s\in N\times S},$$

and a given path for the 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* (11).
2. *The path for* $\{\lambda_{n,s,t}\}_{n,s\in N\times S,\,t\geq 0}$ *satisfies the law of motion of equation* (7).
3. *Population by location-sector* $\{L_{n,s,t}\}_{n,s\in N\times S,\,t\geq 0}$, *satisfies the transition identity* (12).

All equilibrium conditions have an explicit solution. Hence, a unique equilibrium exists and can be written in closed form for any given set of initial conditions and any given path for the exogenous variables.

3.2.4 Existence and uniqueness of a balanced growth path

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

Notice that Equation (7) can be rewritten in matrix form as

$$\vec{\lambda}_{t+1} = \Lambda\vec{\lambda}_t, \tag{13}$$

where $\vec{\lambda}_t$ is a $N\times S$ vector of all scale parameters $\lambda_{n,s,t}$ and Λ is the $(N\times S)^2$ diffusion matrix implied by Equation (7).

From Equation (13), it is immediate to see that, in BGP, $\vec{\lambda}_t$ must be an eigenvector of Λ, with the associated eigenvalue equal to its gross growth rate $1+g^*_\lambda$ (we use upper-stars to denote variables at their BGP value). The Perron-Frobenius theorem guarantees that Λ has a unique positive eigenvector (and corresponding eigenvalue). A sufficient condition for Λ to have only positive entries is that frictions to knowledge diffusion $d_{(m,r)\to(n,s)}$ are positive and finite for each combination of idea origin and destination. This proves the following:[19]

Proposition 1. *Let $0 < d_{(m,r)\to(n,s)} < +\infty$ for all $(m,r),(n,s) \in N \times S$. 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 balanced growth path in which $\{\lambda_{n,s,t}\}_{(n,s)\in N\times S, t\geq 0}$ grow at constant rate g_λ^*.*

3.3 Log-linearized model dynamics

We now study the dynamics of the model by log-linearizing the equilibrium conditions around the BGP. We assume that at time $t-1$ the economy is in a BGP in which the average productivity in each location-sector grows at the same rate and, as a result, the distribution of people across locations is constant. At time t, the economy is hit by technological wave shocks $\{\hat{\alpha}_{r,t}\}_{r\in S}$.[20]

First, consider the dynamics of the scale parameter of the local distribution of productivity, $\lambda_{n,s,t}$. Log-linearizing Equation (7) yields

$$\hat{\lambda}_{n,s,t} = \frac{\theta(\epsilon_{n,s}^*)^\theta}{1+g_\lambda^*} \sum_{m,r} \left(\frac{\lambda_{m,r}}{\lambda_{n,s}}\right)^* \left(\frac{\alpha_r^*}{d_{(m,r)\to(n,s)}}\right)^\theta \hat{\alpha}_{r,t}. \tag{14}$$

Notice that, in BGP, the following relationship holds for each location-sector pair (n,s):

$$g_\lambda^* = (\epsilon_{n,s}^*)^\theta \sum_{m,r} \left(\frac{\lambda_{m,r}}{\lambda_{n,s}}\right)^* \left(\frac{\alpha_r^*}{d_{(m,r)\to(n,s)}}\right)^\theta. \tag{15}$$

Multiplying and dividing the right-hand side of (14) by g_λ^*, and using (8) and (15), we derive the following proposition that links changes in local sectoral productivity to technological wave shocks via the strength of the knowledge diffusion link between the perturbed sector and the receiving location-sector:

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

$$\hat{\lambda}_{n,s,t} = \frac{\theta g_\lambda^*}{1+g_\lambda^*} \sum_{r\in S} \eta_{r\to(n,s)}^* \hat{\alpha}_{r,t}. \tag{16}$$

[19] Huang and Zenou (2020) is another paper that studies the BGP properties of an endogenous growth model with idea diffusion across multiple sectors.

[20] In what follows, we use hats to denote log-deviations from BGP values and stars to denote steady-state values.

The existence of geographical frictions in idea diffusion implies that the reliance on ideas from any given sector r, $\eta^*_{r\to(n,s)}$, depends on the local stock of knowledge in the sector, $\lambda_{n,r}$. From Equation (11) it is also immediate to see that this stock of knowledge is tightly related to the local share of population employed in the same sector.[21] For this reason, Proposition 2 implies that the sensitivity of local productivity to shocks to any given sector is increasing in the weight of the sector in the local economy, and productivity in more diversified locations will be overall less sensitive to technological wave shocks.

Second, consider the population shares $\pi_{n,s,t}$. Combining Equation (11) with the definition $\pi_{n,t} \equiv \sum_{s\in S} \pi_{n,s,t}$ and log-linearizing the resulting expression for any arbitrary deviation of $\lambda_{m,r,t}$ from their BGP values yields

$$\hat{\pi}_{n,t} = \frac{\zeta}{\theta} \sum_{s\in S} \left\{ (1-\pi^*_n)\pi^*_{s|n}\hat{\lambda}_{n,s,t} - \sum_{m\neq n} \pi^*_{m,s}\hat{\lambda}_{m,s,t} \right\}, \tag{17}$$

where $\pi_{s|n}$ denotes the probability of being employed in sector s conditional on living in location n. Equation (17) contains an intuitive condition that controls whether a city grows or shrinks relative to the rest of the economy. A location grows if and only if changes of 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.

To better illustrate the economic mechanism at play, we now consider a simplified version of the model in which knowledge flows across sectors are of second-order importance relative to flows within sectors. Specifically, we impose the following:

Assumption 2. *Frictions to knowledge diffusion across sectors are large enough, so that:*

$$\eta^*_{s\to(n,s)} \approx 1, \qquad \forall s \in S. \tag{18}$$

This approximation allows us to define a measure of exposure to technological waves that only depends on local sectoral shares and aggregate sectoral shocks, resembling the shift-share variable introduced in Equation (1). We can then combine Equations (16) and (17) to derive the following:

Proposition 3. *Under Assumption 2, the percentage change in the population share of location n is proportional to its exposure to technological waves relative to the rest of the economy:*

$$\hat{\pi}_{n,t} = \frac{\zeta g^*_\lambda}{\theta(1+g^*_\lambda)} \underbrace{\sum_{s\in S} \left\{ (1-\pi^*_n)\pi^*_{s|n} - \sum_{m\neq n} \pi^*_{m,s} \right\} \hat{\alpha}_{s,t}}_{RE_{n,t}\,\equiv\, \textit{Relative exposure to technological waves}} . \tag{19}$$

[21] 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 exactly equal to local sectoral share, $\pi^*_{r|n}$.

Proposition 3 rationalizes the reduced-form relationship between exposure to technological waves and population growth documented in Section 2.3. In particular, notice that if the size of any given city is negligible compared to the overall economy,[22] the variation in the right-hand side of Equation (19) is driven entirely by the term $\sum_{s\in S}\pi^*_{s|n}\hat{\alpha}_{s,t}$, which mirrors the measure of exposure in Equation (1).

Proposition 3 also implies that n grows (shrinks) if and only if the average local exposure to the technological wave is larger (smaller) than the average exposure for the rest of the economy:

$$\hat{\pi}_{n,t} > 0 \iff \sum_{s\in S}\pi^*_{s|n}\hat{\alpha}_{s,t} > \sum_{s\in S}\pi^*_{s|-n}\hat{\alpha}_{s,t}, \tag{20}$$

where $\pi_{s|-n}$ is the probability of being employed in sector s conditional on living outside of location n.

3.4 Taking stock

Propositions 2 and 3 show that frictions to knowledge diffusion across geographical areas and technological fields imply rich and heterogeneous effects of technological waves on the evolution of local productivity and on the distribution of population across cities. In the remainder of the paper, we leverage the full structure of the model to quantitatively decompose the effect of technological waves on the dynamics of city growth since 1890 in the United States, to study how local diversification mediates the impact of technological waves, and to speculate on the future evolution of the economic geography under different plausible scenarios of technological trends.

4 Model calibration

In this section, we bring the model to the data to infer the key structural parameters and the unobserved exogenous variables. The model has a recursive structure that allows us to estimate the parameters sequentially by making a limited set of transparent assumptions on how to map the model's objects into data on population, income, and patenting.

In the first step of the calibration, we use model inversion to infer time-invariant amenities u_n, the path of local productivities $\lambda_{n,s,t}$, and aggregate fertility f_t, and simultaneously pin down the structural parameters ζ and θ by matching moments on the dispersion of income and population across cities. We show that the model accurately reproduces the relationship between city size and income despite not being directly targeted. In the second step, we infer

[22] Formally, this requires to impose that $(1-\pi^*_n)\approx 1$ and $\sum_{s\in S,m\neq n}\pi^*_{m,s}\hat{\alpha}_{s,t}$ is approximately equal for all $n \in N$.

the costs of knowledge transmission $d_{(m,r)\to(n,s)}$ by deriving and estimating a gravity equation for idea flows using patent citations data. In the third step, we recover technological wave shocks $\alpha_{s,t}$ and structural residuals $\epsilon_{n,s,t}$ through the law of motion for local productivities.

Throughout the model's calibration, we set 1890 as the starting period and consider the full model dynamics until 2010. We set the model period to 20 years, we let N be the set of 1990 commuting zones that accounted for at least 0.02% of the total population for each decade since 1890, and we define sectors as the 11 IPC class-groups detailed in Appendix Table A.1. The empirical moments used to calibrate the time-invariant parameters correspond to the 1990 observation, for which we have the most recent and complete data on population, income, and patenting.[23]

4.1 Amenities and productivity

As a first step, we jointly calibrate the shape parameters of the Frećhet distributions of utility draws, ζ, and the initial distribution of productivity, θ. Here, we also recover the values of local amenities u_n, and the full path of scale parameters $\lambda_{n,s,t}$ and aggregate fertility f_t.

4.1.1 Productivity distribution

Consider first the scale parameters of the productivity distribution of each location-sector, $\lambda_{n,s,t}$. These objects are at the core of the quantitative analysis: Higher values of $\lambda_{n,s,t}$ imply higher local income, higher ability to attract population, and higher potential to innovate and grow more in the future. In this step of the calibration, we postulate (and later validate) a direct mapping between the stock of patents in a given location-sector and the value of $\lambda_{n,s,t}$. Specifically, we assume that, at any point in time, $\lambda_{n,s,t}$ is equal to a function of current and past patenting:

$$\lambda_{n,s,t} = G_t \times \left[1 + Pat_{n,s,t} + \gamma\, Pat_{n,s,t-1}\right]^{\sigma}, \tag{21}$$

where $Pat_{n,s,t}$ denotes the total number of patents filed at time t in location-sector (n, s) and G_t is a time-variant factor.[24] The parameter γ controls the weight of past patenting on the current stock of knowledge. We set this weight equal to 0.5, which assumes that patents contribute directly to variation in local productivity up to 20 years after filing. The parameter σ represents the elasticity of $\lambda_{n,s,t}$ to the observed stock of patents. This elasticity converts the variation in the local stock of patents into meaningful variation in the average productivity across location-sectors.

[23]Since we assign patents according to their filing year, patents data and citations in the most recent observation (2010) might suffer from truncation issues. Similarly, data on income and population for the 2010 observation are only available from the ACS, that offers a less complete picture than the 1990 census.

[24]We add one to the stock of patents in each sector-city pair to assign a meaningful value to cases in which patenting is zero.

Table 2: **Parameter values and targets**

Parameter	Value	Target	Model	Data
σ	0.22	s.d. log-income (across CZs), 1990	0.19	0.19
θ	2.10	s.d. log-income (overall), 1990	0.64	0.64
ζ	5.90	s.d. log-population (across CZs), 1990	1.07	1.07

Notes: S.D. of log-income for the overall population is taken from Krueger and Perri (2006). S.D. of log-income and log-population across CZs are author's calculations from the NHGIS.

We calibrate σ and θ to jointly match the standard deviation of log-income per capita across cities (in the sample of 373 CZs) and in the overall population in 1990, that are equal to 0.21 and 0.64, respectively.[25] The constant G_t is set to induce an aggregate growth in income per capita of 2% per year.[26]

4.1.2 Amenities, preference draws, and fertility

Consider now local amenities u_n and the shape parameter of the distribution of utility draws ζ. Given any guess for ζ, θ, and $\lambda_{n,s,t}$, we calibrate local amenities to exactly match population by city in the first period (1890).[27] The value of ζ is then calibrated to match the unweighted standard deviation of log-population across cities in 1990. The intuition for the identification is that a higher value of ζ implies lower dispersion in the utility draws among newborn agents, so that differences in the desirability of locations, given by amenities and productivity, are more strongly reflected in migration choices.[28] While we assume time-invariant amenities and do not match population by city in each period, the joint calibration of θ, ζ, and σ guarantees that the equilibrium geography reproduces a realistic dispersion of income and population across locations.

We calibrate the path of fertility f_t to match total population by period in the United States. Notice that, in the absence of moving costs, this is equivalent to assuming that the aggregate increase in population occurs through migration from abroad, fertility, or a combination of the two.

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

[26] 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 period.

[27] We normalize amenities to have a geometric mean of one.

[28] This identification of the dispersion of idiosyncratic preference draws follows a similar intuition as Peters (2019).

4.1.3 Discussion

Table 2 shows the values of θ, ζ, and σ calibrated through this procedure. The corresponding data moments are matched exactly by construction. In Appendix Figure B.3 we show computationally that there are unique values of the three parameters that jointly match those data moments.

There are two key aspects of this calibration strategy that are worth further discussion. First, the mapping of $\lambda_{n,s,t}$ to the stock of patenting (Equation 21) includes a size effect in which larger cities have, other things being equal, higher average productivity. The existence of a correlation between size and productivity is a well-known empirical regularity (see e.g. Glaeser and Gottlieb, 2009) that can emerge as a result of a range of theoretical mechanisms (e.g., sorting, variety, local learning productivity spillovers, higher availability of productive inputs, etc...). While the model is silent on the underlying mechanism behind this correlation (besides the fact that more productive cities will *attract* more population) what is crucial for the quantitative performance of the model is that the resulting elasticity of population with respect to income per capita is empirically accurate. Figure 5 shows a bin-scatter plot of the relationship between log-population and log-income in 1990, both in the model and in the data. Although this correlation is not directly targeted in the calibration, the model captures it closely.[29]

Second, in quantifying the model we assume that residential amenities are time-invariant. This assumption is crucial for the identification of the shape parameter ζ but comes at the cost of not matching population by city exactly after the first period. As we show in Section 5, even without time-varying amenities, the model goes a long way in fitting population growth by city over the last century. But, given a value for ζ, allowing for time-varying residential amenities would be an immediate extension of the model.[30]

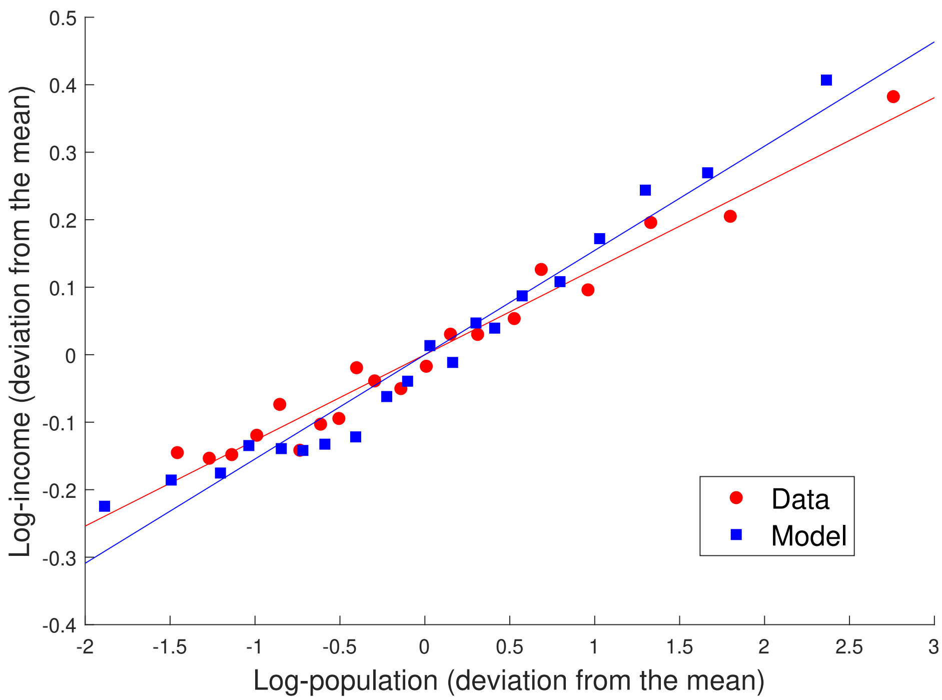
4.2 Gravity equation for knowledge flows

In the second step of the calibration, we derive a simple gravity representation for knowledge flows that we estimate using data on patent citations to recover the parameters controlling knowledge transmission costs ($d_{(m,r)\to(n,s)}$). Specifically, we parametrize frictions to knowledge diffusion as multiplicatively separable between a geographical and a technological component:

[29] The slope of the regression line is equal to 0.155 in the model and 0.127 in the data.

[30] A further extension would be to allow for amenities that combine endogenous and exogenous components. A simple formulation would impose $u_{n,t} = v_{n,t} \times L^{\omega}_{n,t}$, where $v_{n,t}$ is the exogenous component, and ω is the elasticity of amenities to local population, that can account for congestion forces in the case $\omega < 0$. However, the relevant value of ω is likely to change over time and across cities, making the identification particularly challenging.

Figure 5: **Population and Income: Data vs. Model**



Notes: Bin-scatter plot of the relationship between log-population and log-income per capita in the data (red) and the model (blue) in 1990.

$$d_{(m,r)\to(n,s)} = e^{\delta^0_{n,s} + \delta^G \mathbf{1}_{m\neq n} + \delta^K_{r\to s}}, \tag{22}$$

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. We also include a destination fixed effect ($\delta^0_{n,s}$) that we set so that costs are equal to one for flows *within* each location-sector. This normalization is inconsequential for our purposes, since it does not rule out the possibility of systematic differences across receiving location-sectors in their ability to acquire external ideas for innovation, but it bundles those differences with the structural error term ($\epsilon_{n,s,t}$).

Combining Equations (8) and (22) and taking logs on both sides yields

$$\log(\eta^t_{(m,r)\to(n,s)}) = \phi^0_{m,r,t} + \phi^1_{n,s,t} - \theta\delta^G \mathbf{1}_{m\neq n} - \theta\delta^K_{r\to s}, \tag{23}$$

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

Equation (23) illustrates that bilateral citation probabilities $\eta^t_{(m,r)\to(n,s)}$ depend on the composite parameters $\theta\delta^G$ and $\theta\delta^K_{r\to s}$. To recover those composite parameters, we estimate Equation

(23) using data on patent citations across location-sector pairs from the 1990 period (i.e., using patents filed between 1980 and 1999). We compute $\eta^t_{(m,r)\to(n,s)}$ as the share of citations given by patents from (n,s) and directed to patents in (m,r).[31]

Table 3 shows OLS estimates of the composite parameter $\theta\delta^G$. In the baseline specification of Column 1 we replace zero-valued outcomes with the minimum among positive values.[32] Using Column 1 as baseline, in combination with the estimate of θ, we obtain a value for $\delta^G = 2.22$. The coefficient implies highly localized knowledge flows, with the effectiveness of transmission across locations estimated at around 10.9% of the effectiveness of transmission within locations. Notice that, despite the apparent low effectiveness of transmission, the overall weight of ideas from any outside location m may still be large in determining innovation in n, since transmission can happen from *all* the other locations $m \neq n$. Column 2 shows the same regression when only positive values of $\eta^t_{(m,r)\to(n,s)}$ are used in the estimation. The estimate still reveals highly localized knowledge flows, but the coefficient declines in absolute value, suggesting that, as expected, zero values are concentrated among pairs of different locations.

The same regression also delivers a full set of bilateral transmission costs across sectors ($\delta^K_{r\to s}$), that we show in a heatmap in Figure B.2 in the Appendix. As expected, these costs are estimated to be lower within sectors (along the diagonal of the heatmap), although all pairs of sectors display some degree of knowledge exchange that, in some cases, is far from negligible, such as in the cluster of class-groups G1 ("Physics") and H1 ("Electricity").

4.3 Technological waves and structural residuals

In the third step of the calibration, we use the estimates of θ and $\delta^K_{r\to s}$ and the values of $\lambda_{n,s,t}$ in combination with the law of motion (7) 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 (7) to recover the full set of structural residuals. By construction, this step rationalizes the path of $\lambda_{n,s,t}$ for any initial guess of $\{\alpha_{s,t}\}_{s\in S}$. Hence, to complete the identification, we need to impose an additional condition. We assume that the average growth in productivity for each sector in the long-run is fully explained by technological waves and their interaction with the endogenous process of knowledge creation and diffusion. Idiosyncratic residuals explain the remaining variation in productivity growth across locations. Specifically,

[31] 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. Every citing patent in our regression has a total weight of one. In other words, every observation is weighted by the inverse of the total number of citations given by (n,s). In order to include all patents in the estimation, we further assume that every grant gives at least one citation to its own location-sector.

[32] Notice that $\eta^t_{(m,r)\to(n,s)}$ cannot be written as a count divided by an exposure variable, so Equation (23) cannot be estimated via Poisson Pseudo-Maximum Likelihood.

Table 3: **Gravity equation for knowledge flows**

	Log share of citations	
	(1)	(2)
Origin CZ $\neq$ Destination CZ	-4.654***	-2.270***
	(.048)	(.021)
Origin location-sector FE	yes	yes
Destination location-sector FE	yes	yes
Origin-Destination sector FE	yes	yes
# Obs.	16,834,609	1,255,496
R^2	0.30	0.67
Zero values	Set to min	No

Notes: OLS estimates. Observations are all the combinations of pairs of location-sectors. The dependent variable is the logarithm of the share of citations given by each destination location-sector to each origin location-sector, where each cting patent is given a weight of one. All patents are assigned one citation to their own location-sector. Standard errors clustered at the destination location-sector level in parenthesis. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

we impose adjusted structural residuals to have a weighted average of 1 for each sector and time period:

$$\mathbb{E}\left[\epsilon^{\theta}_{n,s,t}\right] = 1, \qquad \forall s \in S,\, t \geq 0. \tag{24}$$

It is important to emphasize that we do not make any assumption on the nature and properties of the structural residuals, including on whether they are stochastic or deterministic, what is their spatial and temporal correlation, and whether they are systematically correlated with the other terms in Equation (7). We discuss this point in detail in the next section.

5 Quantitative results

In this section, we explore the ability of the model to account for the evolution of the U.S. economic geography in the period 1890-2010. We start by showing that the interaction between technological waves and the endogenous mechanism of knowledge creation and diffusion has a sizeable impact on the long-run growth rate of cities. We examine the importance of this mechanism in accounting for the two most striking episodes of technological and geographical transformation of the last century: The extraordinary rise of manufacturing-intensive cities in the early decades of the 20th century and their later decline as cities specialized in knowledge-intensive sectors gained prominence. These two episodes illustrate that technological waves, interacting with frictions to knowledge diffusion, can explain at the same time path dependence and reversal of fortune in the growth trajectory of cities. We then use the model to quantitatively assess the role of diversification in mediating the effect of technological waves, and to predict how the economic geography of the U.S. might evolve in the coming decades in response to possible changes in the technological environment.

5.1 The impact of technological waves on population growth

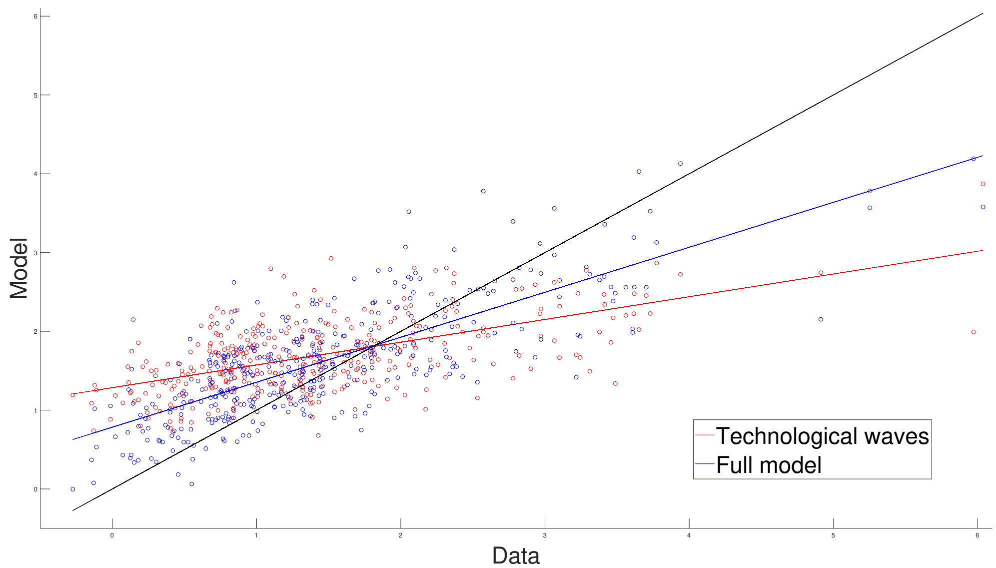
Figure 6 shows the performance of the model in accounting for the variation in city growth since 1890. The graph plots the 1890-2010 difference in the log-population of the sample of commuting zones in the data (horizontal axis) and the model (vertical axis). Since local amenities are fixed at the initial period, the full model does not replicate the data exactly. Hence, the 45-degree (black) line can be taken as a benchmark for performance.

The blue line and circles correspond to the complete model with both aggregate technological wave shocks ($\alpha_{s,t}$) and structural residuals ($\epsilon_{n,s,t}$).[33] The slope of the regression line is 0.57 and the correlation between the predicted and actual values is 77%. The red line and circles correspond to the case where we initialize the equilibrium by feeding the full set of shocks in the first period of the dynamics. We then keep structural residuals constant and only input technological wave shocks. In this case, we let the path for $\lambda_{n,s,t}$ to be determined by the endogenous law of motion in Equation (7), which reflects the interaction between the state variables in 1910 and the gradual unfolding of technological waves over time. The predictive power of the model declines but remains significant: The slope of the line is 0.29, while the correlation is 58%. This correlation can be interpreted as the contribution of technological waves, interacting with the endogenous process of innovation and knowledge diffusion, in explaining the variation in population growth over the last century.

Figure 7 shows further evidence of the quantitative importance of technological wave shocks in shaping the economic geography. We show bin-scatter plots of the relationship between relative exposure to technological waves and population growth, as predicted by Proposition 3, in which variables are built using the calibrated model objects. The blue line corresponds to the complete model, in which we feed the full path of technological waves and structural residuals. The red line corresponds to the counterfactual dynamics, in which we feed the path of technological waves but keep structural residuals fixed at their 1910 values. The prediction of Proposition 3 is confirmed in both scenarios. The slope of the relationship in the full

[33] This corresponds to feeding the true path of $\lambda_{n,s,t}$ as inferred via Equation (21).

Figure 6: **Population growth: Data vs. Model**



Notes: The graph plots the 1890-2010 log-difference in population across CZs in the data (horizontal axis) and the model (vertical axis). The black line corresponds to the 45-degree line (perfect fit). The blue line shows the full model (with structural residuals and technological waves). The red line shows the model in which structural residuals are kept constant to 1910.

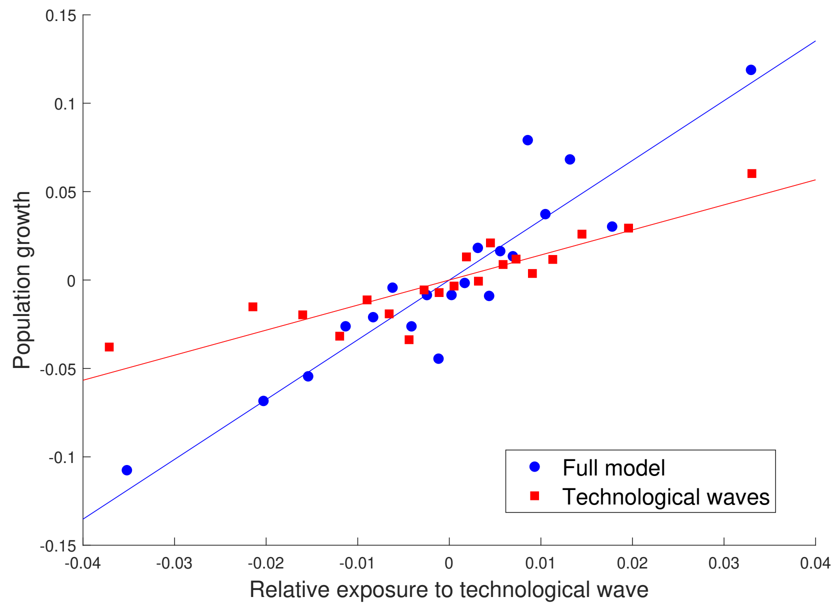
model is larger than the one in the counterfactual, suggesting that, on average, residual factors amplify, rather than attenuate, the impact of technological waves on population growth. These amplifying factors may include agglomeration economies and positive externalities in local investment and residential amenities. Note that, although the residual amplifies the effect on average, the model allows for cases in which the residual *attenuates* the effect of technological waves, for instance through progressive taxation and redistributive policies, migration frictions, land-use restrictions, congestion, and other convex costs of local inputs.

5.2 The rise of manufacturing-intensive cities

We now look specifically at how the calibrated model can account for two of the most striking episodes of transformation of the economic geography of the U.S. in the last century: The extraordinary rise of manufacturing-intensive cities in the early 20th century, followed by their decline and simultaneous rise of knowledge-intensive urban areas.

Figure 8 shows, for a selected subset of cities in the U.S., population growth between 1910 and 1950 under the full model, that includes technological waves and structural residuals (blue line), and the model in which we only provide the path for technological waves but keep structural residuals constant at their 1910 value (red line). We normalize data so that the

Figure 7: **Relative exposure to the technological wave and population growth**

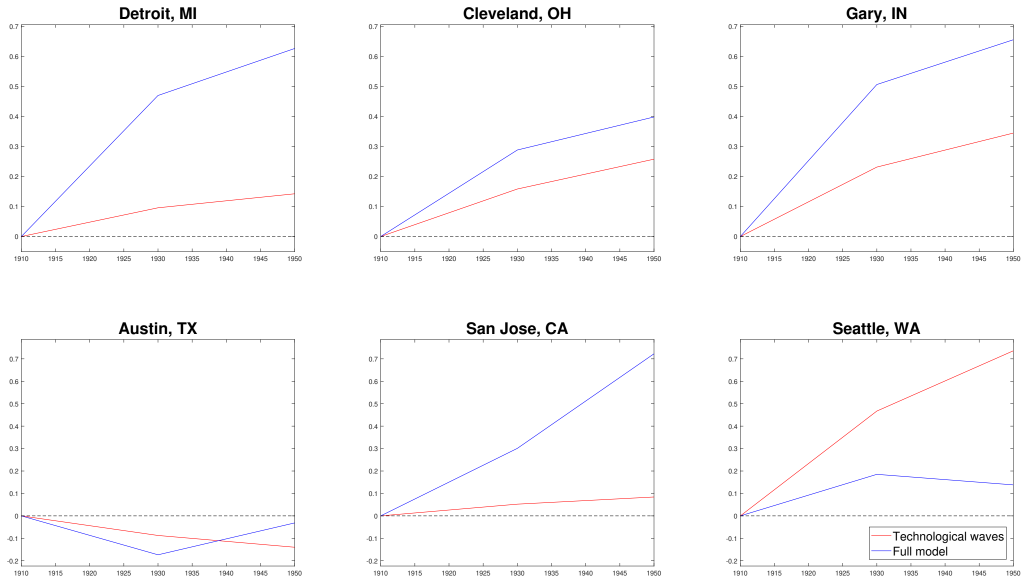


Notes: Bin-scatter plot of the relationship between relative exposure to the technological wave and population growth in the full model (blue) and in the model with only technological wave shocks (red). Variables are residualized with respect to decade fixed effects.

horizontal axis corresponds to a scenario in which all cities grow at the same rate, as dictated by aggregate fertility. The plots show that Detroit and Cleveland, the largest urban areas of what would later be known as the Rust Belt, not only were favorably exposed to the technological wave in the early part of the century, but were also subject to forces – captured by the structural residuals – that significantly amplified the effect of their exposure. According to the model, the 1910 exposure to the technological wave resulted in higher population of 14% in Detroit and 26% in Cleveland in 1950 (red line). These effects are significantly larger in the model with the full set of shocks (blue line) that accounts for factors that evolve endogenously to the local response to the technological wave (such as agglomeration externalities) and exogenous forces that affect population growth by boosting local innovation and productivity (such as investment in infrastructure uncorrelated with other local disturbances). Although a systematic exploration of those factors goes beyond the scope of this paper, the model allows us to infer, for each episode, whether those forces amplified or dampened on net the effect of the technological wave. In this scenario, Detroit and Cleveland are 63% and 40% larger in 1950 compared to the baseline, respectively. A similar path was experienced by Gary, IN, another major center of the Rust Belt that experienced fast growth in the early 20th century driven by the expansion of the local steel industry.

During the same period, cities did not uniformly benefit from this transformation in the

Figure 8: **Growth Decomposition: 1910-1950**



Notes: The blue and red lines show log-population in deviation from a trajectory of constant population growth across cities, as dictated by aggregate fertility. The blue line corresponds to the full model (with structural residuals and technological wave shocks). The red line corresponds to a model in which the structural residuals are kept constant to 1910.

technological landscape. As described in Section 2.2, the commuting zone of Austin lost population. The red line shows that part of this decline was due to an unfavorable exposure to the technological wave. However, the blue line shows that, in the period 1910-1930, external forces were even more penalizing. The Texas Oil Boom created opportunities in the other areas of Texas, further depressing population growth in Austin. In this period, San Jose was positively, albeit weakly, exposed to the technological wave, but residual factors, such as a general expansion of the West, strongly contributed its population growth. Finally, the commuting zone of Seattle was positively affected by the technology cycle, as reflected by the rising importance of technologies related to shipbuilding first and aviation later between WWI and WWII. However, in this case, contrary to experiences of Detroit and Cleveland, the model records a negative contribution from residual factors.

5.3 The emergence of modern knowledge hubs

The experiences of Detroit, Cleveland, and Gary in the first half of the 20th century were not isolated cases. Several other cities that specialized in heavy manufacturing and were mostly concentrated in what is now known as the Rust Belt witnessed exceptional growth in population.

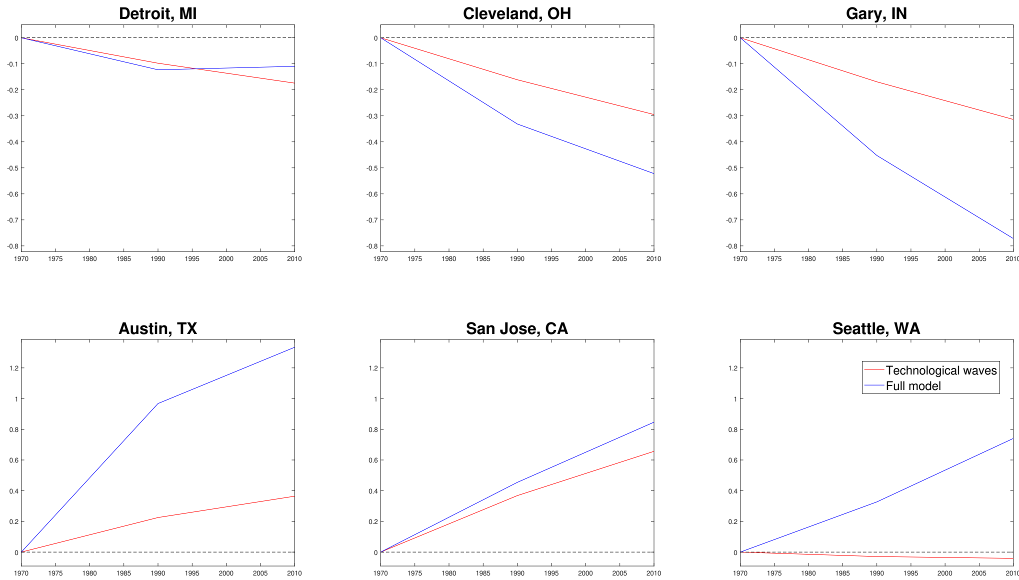
The model suggests that part of this growth can be explained by the availability of local ideas in fields that were complementary to the prevailing technological wave. We now explore whether the same factors that led to the remarkable growth of manufacturing-intensive cities contributed to their later decline to the benefit of emerging knowledge hubs specialized in information technologies.

The top panels of Figure 9 track the response of the three manufacturing-intensive cities of Figure 8 in the later part of the century. All those commuting zones experience a negative direct impact of the technological wave on city population (red line) of roughly 18% in Detroit, 29% in Cleveland, and 31% in Gary. The residual factors in the evolution of productivity led to more dramatic declines of 52% and 77% in Cleveland and Gary, respectively, but not in Detroit, where the decline in population in the full model is attenuated (11%). The reason why the structural residuals impose a more severe loss in the commuting zones of Cleveland and Gary compared to Detroit is interesting and worth further investigation. One possible explanation is that the policy response to the decline of the automotive industry compressed the amplification mechanisms in Detroit but not in Cleveland and Gary.

Throughout the same decades, a handful of cities emerged as modern leading technological hubs. The commuting zones of Austin, TX and San Jose, CA are archetypal examples of this expansion. The model suggests that population in Austin and San Jose increased, relative to the baseline, by 36% and 65%, respectively, as a direct effect of the technological wave interacted with their local characteristics in 1970. However, the amplification effect coming from the structural residuals is significantly larger for Austin than it is for San Jose. Why does the contribution of structural residuals vary so much in these two cases? While an definitive answer to this question is beyond our scope, a candidate explanation can be found in the different constraints imposed by local taxation and land-use regulation that characterize those commuting zones. This hypothesis is in line with the evidence in recent studies, such as Glaeser and Ward (2009) and Hsieh and Moretti (2019), that document the consequences of land-use restrictions on the misallocation of people across U.S. cities.

Finally, the commuting zone of Seattle appears to be weakly but negatively affected by the technological wave, and to receive instead a positive contribution from structural residuals. This finding is in line with the fact that most of the recent growth in the IT sector in Seattle is a consequence of local events that happened after 1970 (such as the relocation of Microsoft to Bellevue in 1979 and the establishment of Amazon in 1994). In fact, we do find a positive direct effect of the technological wave if we consider the model with the full set of shocks until 1990, and only provide the technological wave shocks in 2010.

Figure 9: **Growth Decomposition: 1970-2010**



Notes: The blue and red lines show log-population in deviation from a trajectory of constant population growth across cities, as dictated by aggregate fertility. The blue line corresponds to the full model (with structural residuals and technological wave shocks). The red line corresponds to a model in which the structural residuals are kept constant to 1970.

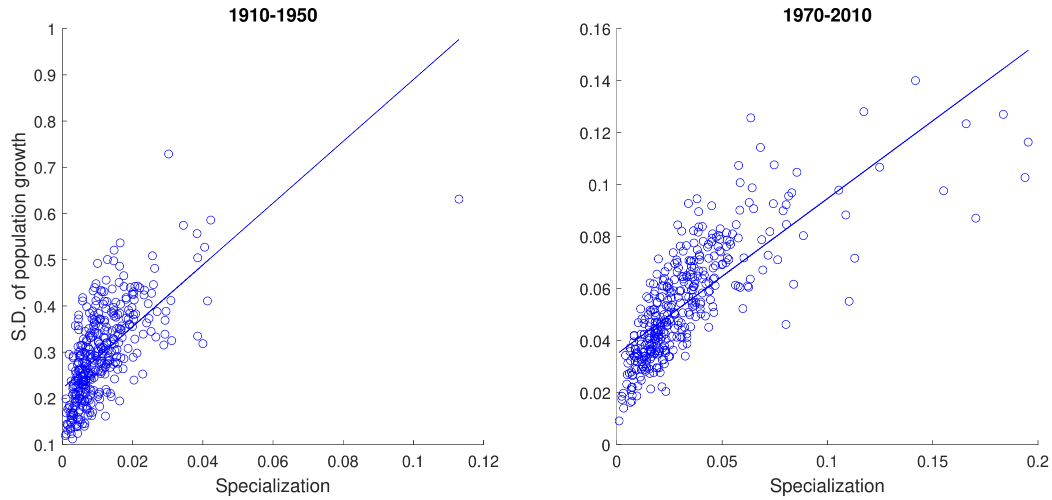
5.4 Diversification and resilience to technological waves

The process of innovation through frictional idea diffusion implies a role for local diversification in determining cities' resilience to changes in the technological landscape. There are two channels that make the growth trajectory of diversified cities less sensitive to technological wave shocks, reflecting respectively the existence of frictions to knowledge diffusion across technological fields and geographical areas.

First, the existence of frictions to knowledge diffusion *across technological fields* implies that the path of productivity of any given sector is disproportionately driven by technological wave shocks to the same sector. This also implies that in diversified cities, whose sectoral composition is dispersed across multiple sectors, average productivity will be less volatile, since negative shocks to some sectors are likely to be compensated by positive shocks to other sectors.

Second, the existence of frictions to knowledge diffusion *across geographical areas* implies that the reliance of each location-sector on ideas from any given field is an increasing function of the *local* availability of ideas from that field. For this reason, in more diversified cities, innovation in any sector relies on ideas from a broader set of fields. This also implies that the path of productivity of any local sector is less sensitive to technological wave shocks to specific

Figure 10: **Specialization and growth volatility**



Notes: Relationship between the measure of specialization in Equation (25) and the standard deviation of the log change in population across 500 simulations.

sectors.

To verify that more diversified cities experience less volatile growth trajectories in response to technological wave shocks, we perform simulations in which we randomly reshuffle the path of shocks experienced by each sector, $\{\hat{\alpha}_{r,t}\}_{r\in S}$, and compute the corresponding equilibrium for the economy. We then correlate the standard deviation of population growth across all the simulations with a measure of local specialization. Specifically, we define local specialization as the Euclidean distance between the local and nationwide vectors of sectoral shares:

$$Spec_{n,t} = \sum_{s\in S}\left(\pi_{s|n,t} - \sum_{m\in N}\pi_{m,s,t}\right)^2. \tag{25}$$

As we show in Proposition C.1, under Assumption 2 and intuitive conditions on the distribution of shocks (Assumption C.1), the theory predicts that this measure of specialization is approximately proportional to the variance of local population growth.

Figure 10 plots the correlation between $Spec_{n,t}$ and the standard deviation of population growth across 500 simulations in the early period (1910-1950, left panel), and the late period (1970-2010, right panel) separately. The figure shows a consistently positive correlation, indicating that diversification makes the growth trajectory of cities less sensitive to technological wave shocks. The correlation is strong (the R^2 of the regressions are equal to 0.40 and 0.59, respectively), and the size of the effect is meaningful. A one-standard deviation increase in specialization leads to an increase in the standard deviation of 63.7% of a standard deviation in the early period, and of 77.0% of a standard deviation in the late period.

5.5 The U.S. economic geography under future technological waves

The quantitative model can be used to predict the evolution of the U.S. economic geography in the coming decades in response to transformations in the innovation landscape. In this Section, we propose plausible scenarios for future technological waves and look at which commuting zones will be most positively and negatively affected by those changes. In particular, we project population growth across cities until 2050 under different assumptions on the evolution of the importance of different sectors ($\alpha_{s,t}$), and compare the outcome with a baseline in which the importance for all sectors is kept constant to its 2010 values.

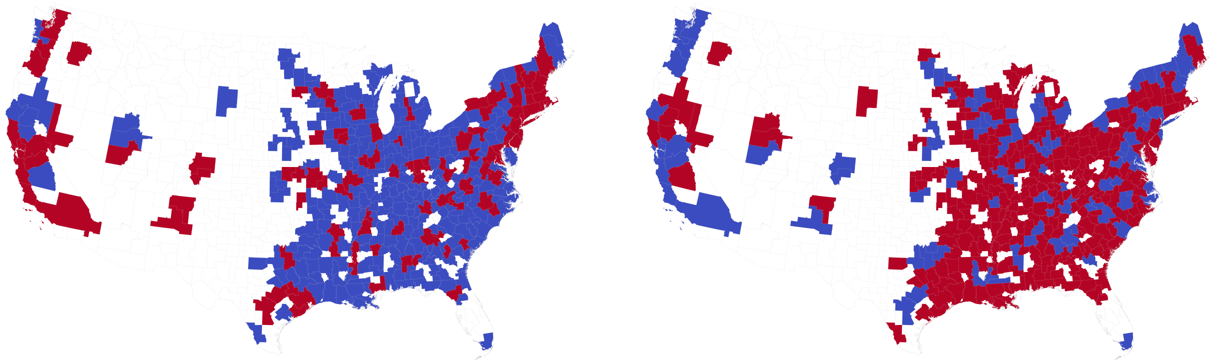
In the first scenario, we assume that sector B2 (”Transporting”) experiences a technological wave shock of magnitude +8.6% (equal to twice its standard deviation throughout the sample period) as new advances in transit technologies and autonomous vehicles induce innovation in transportation to return to a pivotal role. The left map in Figure 11 visually illustrates the results. Commuting zones in blue (red) experience a net gain (loss) of population compared to the baseline. Given the current state variables, cities in the Rust Belt are the areas that are best positioned to take advantage from this transformation. Detroit would experience a 4.6% increase in population compared to the baseline. Other urban centers of manufacturing related to (but not specialized in) transportation, would benefit, too, albeit to a lesser extent. For example, Cleveland and Gary would increase population by 0.3% and 1.5%, respectively. The knowledge hubs of Austin (-2.2%), San Jose (-2.8%), and Seattle (-1.4%), would all experience a relative loss of population.

An alternative way of modelling transportation-related technologies gaining prominence in the innovation landscape is to assume that ideas from B1 become more relevant for innovation in either G1 (“Physics”) or H1 (“Electricity”), and vice versa. An immediate example of the increasing inter-dependence of those sectors is the gradual integration of IT components in electrical and autonomous vehicles. We model this strengthening connection as a drop in the cost of knowledge transmission ($\delta_{s\to r}$) by assuming a 10% decline in composite knowledge frictions ($d^{\theta}_{(m,r)\to(n,s)}$) from (to) B4 to (from) both G1 and H1.[34] In this case, we keep sectoral importance ($\alpha_{s,t}$) at its 2010 value. The right map in Figure 11 displays the results. In this case, both Detroit (+1.5%) and the prominent knowledge hubs of Austin (+2.0%), San Jose (+1.6%), and Seattle (+1.9%) experience a relative gain in population. The decline in the transmission costs between sectors effectively provides better innovation possibilities for cities specialized in either of the affected fields.

In the second scenario, we simulate a large positive technological wave shock to sector A3 (“Health; Life-Saving; Amusement”, that includes the bulk of innovation related to pharma-

[34] While the assumption of a proportional 10% decline is arbitrary, this choice only affects the magnitude of the results but does not alter the qualitative patterns.

Figure 11: **Counterfactuals: Autonomous vehicles**

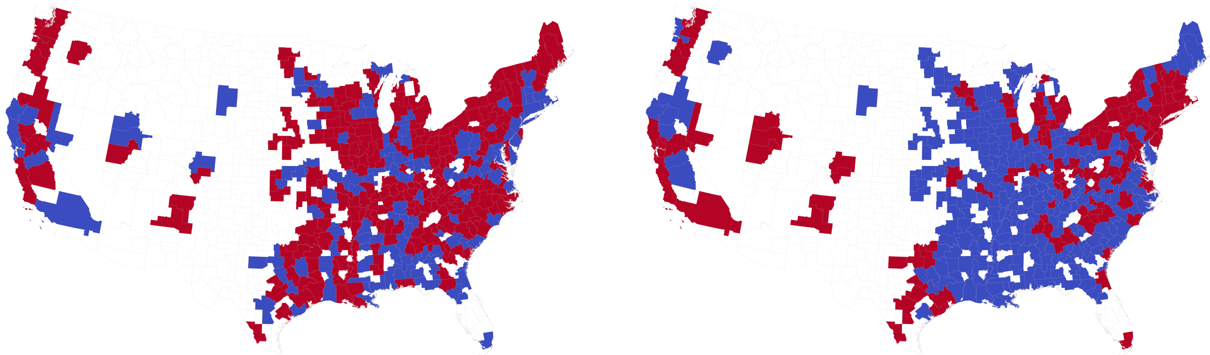


Notes: The map shows log-population in 2050 after a technological wave shock to B2 of magnitude +8.6% (left map), and a 10% decline in composite knowledge frictions ($d^{\theta}_{(m,r)\rightarrow(n,s)}$) from (to) B2 to (from) both G1 and H1 (right map), in deviation from a status quo in which $\alpha_{s,t}$ are kept at their 2010 values. Blue CZs correspond to a net population gain, red CZs correspond to a net population loss.

ceuticals and medical sciences) possibly in response to new challenges in global health such as the COVID-19 pandemic. We input a shock of magnitude 23.8%, equal to twice its standard deviation throughout the sample. The results are depicted in the left map of Figure 12. The counterfactual suggests that major commuting zones in the North-East, such as Boston (+8.7%) and Providence (+19.1%), and in California, such as Los Angeles (+7.0%) and San Francisco-Oakland (+3.6%) would experience a net inflow of population, at the expense of IT clusters such of Austin (-11.1%), San Jose (-5.7%) and Seattle (-5.2%).

In the third scenario, we assume that sector A1 ("Agriculture"), regains centrality by experiencing a 15.4% technological wave shock (twice its standard deviation throughout the sample). This is a plausible scenario that can emerge as a result of tightening regulatory constraints and shifting demand towards sustainable farming, possibly in response to global challenges such as climate change. Results are in the right map of Figure 12. Under this scenario, the economic geography of the U.S. experiences a pronounced shift away from the East and West coast and the Rust Belt, towards the Central States. Among the major commuting zones, Des Moines (IA) receives the highest net gain (+13.4%). This scenario would represent a significant convergence force in relative population across commuting zones: A regression of log-population in 2010 with the log-deviation from the baseline in 2050 delivers a coefficient of -1.0%, implying that population would mostly relocate away from larger commuting zones and towards less-populated ones.

Figure 12: **Counterfactuals: Pharmaceuticals and Agriculture**



Notes: The map shows log-population in 2050 after a technological wave shock to A3 of magnitude +23.8% (left map), and after a technological wave shock to A1 of magnitude +15.4% (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, red CZs correspond to a net population loss.

6 Conclusions

The economic geography of countries is characterized by rich and uneven dynamics, alternating persistence to occasional reversal of fortunes. Some cities remain large and important throughout long time spans, while others experience episodes of sharp growth and decline. In this paper, we explore and quantify the hypothesis that these rich dynamics result in part from cities' patterns of specialization across sectors, coupled with the continuous evolution of the technological landscape. We develop a parsimonious framework that combines elements from quantitative spatial equilibrium models and theories of endogenous growth through innovation and idea diffusion. The model remains tractable for any arbitrary number of sectors, locations, and time periods – making it suitable for quantitative analysis – and delivers a wide range of predictions on how the economic geography of countries responds to changes in the technological environment. The reduced-form and structural analysis support the idea that the interaction of frictional knowledge diffusion with technological waves played a major role in shaping the evolution of the U.S. economic geography in the last century, accounting for 58% of the variation in population growth across cities between 1890 and 2010. The model can account for the rise in manufacturing-intensive cities in the Rust Belt, driven by the increase in the centrality of transportation technologies, and the recent emergence of modern knowledge hubs, driven by the increase in the centrality of fields related to physics and electricity. We use the model to speculate on how the U.S. economic geography will evolve under different technological scenarios, such as a come back of transportation and agriculture and a further rise in the centrality of medical sciences.

While we consistently find a large impact of technological waves through frictional knowledge diffusion, there are some moderating channels from which the model abstracts but could be embedded without a prohibitive loss of tractability. Extensions can include trade and migration frictions across locations, production activities that are geographically separated from innovation, the inclusion of local non-tradable goods, and partially endogenous residential amenities and structural residuals. In particular, the quantitative results suggest that residual factors contribute significantly to the dynamics of local innovation and to the variation in city growth. The framework allows us to isolate the direct effect of the technological wave via innovation and knowledge diffusion, and does not require to make specific assumptions on the nature of this residual. A possible way to endogenize this error term is to allow innovators to exert effort to improve their ideas, in the spirit of an endogenous growth theory with expanding varieties (as in Jones, 2005). An alternative route to unpack the residual term is to account for the granularity of the locational choices of individual firms. Events such as Microsoft's relocation to the Seattle area, or Amazon's selection of a site for its second headquarters, can have a major impact in shaping the destiny of cities. Other endogenous forces that enter the residual include congestion, pecuniary externalities on local assets, and the response of policy to local shocks. Understanding how these factors contribute to amplifying or dampening the effects of technological waves is the next step in our agenda.

References

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.

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 (2018): "Geography and path dependence," Tech. rep.

Andrews, M. (2019): "Historical Patent Data: A Primer and Some Pitfalls," mimeo, available at: https://sites.google.com/site/michaeljeffreyandrews/research.

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.

Babina, T., A. Bernstein, and F. Mezzanotti (2020): "Crisis Innovation," *Available at SSRN*.

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?" .

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

Berkes, E. and P. Nencka (2020): "Knowledge Access: The Effects of Carnegie Libraries on Innovation," Available at SSRN: https://ssrn.com/abstract=3629299.

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

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 Jr (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.

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

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.

Clemens, J. and P. Rogers (2020): "Demand Shocks, Procurement Policies, and the Nature of Medical Innovation: Evidence from Wartime Prosthetic Device Patents," Tech. rep., National Bureau of Economic Research.

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.

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 (2019): "Spatial Structural Change," mimeo, available at: https://mipeters.weebly.com/research.html.

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 B. A. Ward (2009): "The causes and consequences of land use regulation: Evidence from Greater Boston," *Journal of Urban Economics*, 65, 265–278.

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.

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.

Hsieh, C.-T. and E. Moretti (2019): "Housing constraints and spatial misallocation," *American Economic Journal: Macroeconomics*, 11, 1–39.

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.

Jones, C. I. (2005): "Growth and ideas," in *Handbook of economic growth*, Elsevier, vol. 1, 1063–1111.

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

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.

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.

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

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

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.

Peters, M. (2019): "Market Size and Spatial Growth–Evidence from Germanys Post-War Population Expulsions," *Unpublished manuscript*.

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

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.

A Additional tables

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	2,984	353,213	897,673	225	1.79e+07
Log-population	2,984	11.93	1.16	5.42	16.70
Population growth	2,611	.254	.356	-.606	4.29
Total patents (CZ)	2,891	1,257.8	5,086.4	0	77,998
Log-total patents	2,878	4.88	2.10	0	11.26
Patents per capita (CZ)	2,891	.001	0.002	0	0.035
Log-patents per capita	2,878	-7.13	1.28	-11.35	-3.35

Notes: Summary statistics are unweighted and refer to the period 1870-2010, with the exception of Population growth, that refer to the period 1890-2010.

Table A.3: **Technological waves and city growth: Earlier vs. later samples**

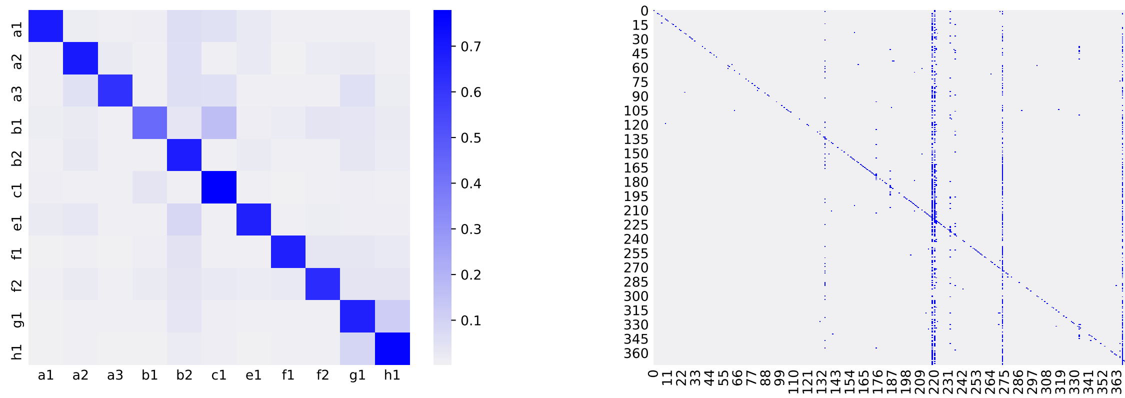
	Growth rate of population					
	1890-1930			1950-2010		
	(1)	(2)	(3)	(4)	(5)	(6)
Relative exposure to tech. wave	.665**	.566**	.572**	.323***	.338***	.269***
	(.231)	(.241)	(.245)	(.076)	(.068)	(.064)
Lagged log-population	-.216***	-.196**	-.194**	.031**	.036***	.020**
	(.075)	(.084)	(..083)	(.012)	(.009)	(.009)
Human capital (ranking)			-.073			.170***
			(.150)			(.044)
Fixed effects	T	CD×T	CD×T	T	CD×T	CD×T
# Obs.	1,119	1,119	1,119		1,492	1,492
R^2	0.34	0.44	0.44	0.06	0.33	0.36

Notes: CZ level regression, 1890-1930 (columns 1-3) and 1950-2010 (columns 4-6). Dependent variable defined as growth rate of population over 20 years. Observations are weighted by the share of population at the beginning of the period. “T” denotes time fixed effects, and “CD×T” denotes Census Division-time fixed effects.. Standard errors clustered at the CZ level in parenthesis. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

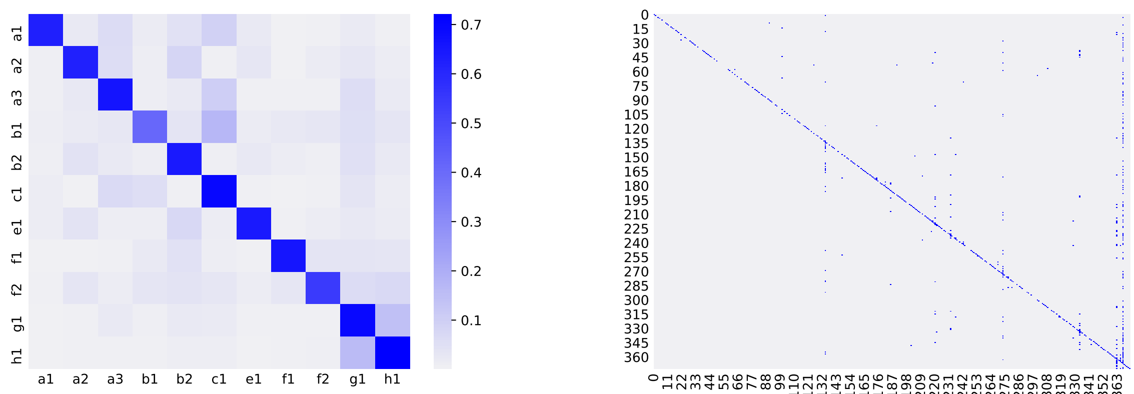
B Additional figures

Figure B.1: **Patent citations across fields (left) and locations (right), by period**

1950-1979

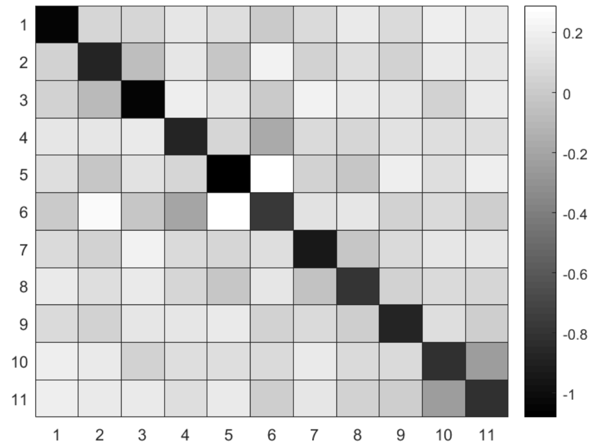


1980 onwards



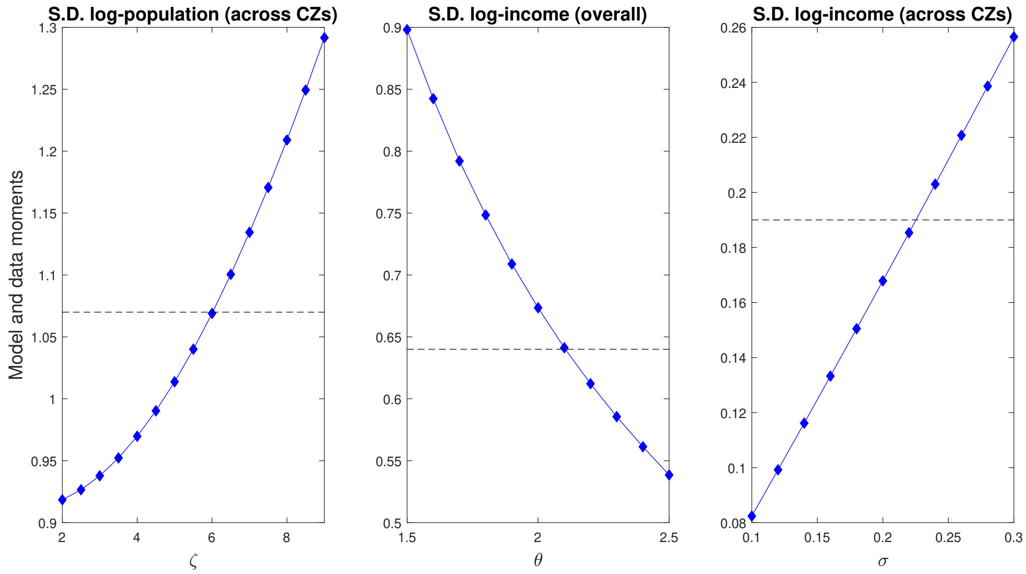
Notes: Left panels: Probability that patents from the class-group on the vertical axis cite patents from the class-group on the horizontal axis. Right panels: Probability that patents from the commuting zone on the vertical axis cite patents from the commuting zone on the horizontal axis.

Figure B.2: **Knowledge transmission costs across sectors**



Notes: Estimated (OLS) coefficients $\delta^K_{r \to s}$, from regression of Table 3, column 1. Observations are all the combinations of pairs of location-sectors. The dependent variable is the logarithm of 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. All patents are assigned one citation to their own location-sector. Rows correspond to citing (idea destination) sectors. Columns correspond to cited (idea origin) sectors. Number of observations: 16,834,609. R^2: 0.31.

Figure B.3: **Identification of ζ, θ, and σ**



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 fixed at their calibrated values. Both model and data refer to the 1990 observation.

C Derivations

To rationalize the measure of local specialization in Equation (25), we impose the following assumption on the distribution of technological wave shocks:

Assumption C.1. *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 C.1 in combination with Assumption 2 we derive the following theoretical result, that links the volatility of local population growth to the local degree of specialization:

Proposition C.1. *Under Assumptions 2 and C.1, the variance percentage change in the population share of location* n *is proportional to the local degree of specialization:*

$$Var(\hat{\pi}_{n,t}) \propto \underbrace{(1-\pi_n^*)^2 \sum_{s\in S} \left(\pi_{s|n}^* - \pi_{s|-n}^*\right)^2}_{Spec_n^* \equiv \, Specialization}. \tag{26}$$

Proof. Consider the definition of "Relative exposure to technological waves" introduced in Proposition 3:

$$RE_{n,t} \equiv \sum_{s\in S} \left\{ (1-\pi_n^*)\pi_{s|n}^* - \sum_{m\neq n} \pi_{m,s}^* \right\} \hat{\alpha}_{s,t}. \tag{27}$$

Factoring out $(1-\pi_n^*)$, and realizing that $\pi_{s|-n}^* \equiv \sum_{m\neq n} \frac{\pi_{m,s}^*}{1-\pi_n^*}$, we can rewrite $RE_{n,t}$ as

$$RE_{n,t} \equiv (1-\pi_n^*) \sum_{s\in S} \left(\pi_{s|n}^* - \pi_{s|-n}^*\right) \hat{\alpha}_{s,t}.$$

Under Assumption C.1, the technological wave shocks $\alpha_{s,t}$ have zero covariance and common variance V. Hence, the variance of $RE_{n,t}$ is equal to

$$Var(RE_{n,t}) = V \times (1-\pi_n^*)^2 \sum_{s\in S} \left(\pi_{s|n}^* - \pi_{s|-n}^*\right)^2$$

Defining "Specialization" as

$$Spec_n^* \equiv (1-\pi_n^*)^2 \sum_{s\in S} \left(\pi_{s|n}^* - \pi_{s|-n}^*\right)^2,$$

and using Proposition 3 to have $\hat{\pi}_{n,t} \propto RE_{n,t}$ yields

$$Var(\hat{\pi}_{n,t}) \propto Spec_n^*.$$

$\square$

Notice that if all cities are negligible in size compared to the overall economy, the measure in Equation (26) approximates the measure of specialization in Equation (25).