Income Segregation and the Rise of the Knowledge Economy*

Enrico Berkes† Ruben Gaetani‡

8th September 2021

Abstract

We analyze the effect of the rise of knowledge-intensive activities on spatial inequality within U.S. cities. We use the predetermined network of patent citations to instrument for local trends in innovation. Between 1990 and 2010, a one standard deviation increase in patenting growth increases income segregation by 0.65 Gini points, corresponding to 30.6% of a standard deviation in the change in income segregation. The effect is mainly driven by the sorting of residents in terms of income, occupation, and education. Local shocks to innovation induce a clustering of knowledge-intensive jobs and residents, amplified by the response of rents and amenities.

Keywords: Cities, Innovation, Economic Segregation.

JEL Classification: O15, O33, R11.

*We thank Treb Allen, Davide Coluccia, Matthias Doepke, Richard Florida, Ben Jones, Lorenz Kueng, Marti Mestieri, Peter Nencka, Matthew Notowidigdo, Bruce Weinberg, our discussants Ed Glaeser, Henry Overman and Jeffrey Zabel, as well as seminar participants at Northwestern University, University of Toronto, McMaster University, Ryerson University, EIEF, Bocconi, CREI, INSEAD, Toulouse School of Economics, Ohio State University, Carnegie Mellon University, Chicago Fed, Auburn University, 2016 Conference of Swiss Economists Abroad (Bern), 2017 European Meeting of the Urban Economic Association (Copenhagen), 2017 North American Meeting of the Urban Economic Association (Vancouver), NBER Trade and Geography Conference (Cambridge), 2018 North American Summer Meeting of the Econometric Society (Davis), 2018 Meeting of the Society of Economic Dynamics (Mexico City), NBER Summer Institute Innovation (Cambridge), Duke Fuqua Strategy Conference, Conference on Urban and Regional Economics (Philadelphia Fed), and attendees at the 2017 Workshop of the Kauffman Foundation (Chicago) for their helpful comments. We thank Fernando Crupi for providing excellent research assistance and Catherine Wieczorek for providing help with the graphics. This research was funded in part by the Ewing Marion Kauffman Foundation. The contents of the paper are solely our responsibility.

†Ohio State University, berkes.8@osu.edu.
‡University of Toronto, ruben.gaetani@utoronto.ca.
1 Introduction

Over the past 40 years, the economic activities that rely on non-manual, non-routine technical skills, scientific knowledge, and intellectual creativity have become the main engine of economic prosperity in advanced countries (Powell and Snellman, 2004). Since 1975, the share of value added generated by knowledge-intensive sectors in the United States has increased by almost 15 percentage points, and the number of patents per capita issued by the United States Patent and Trademark Office (USPTO) has doubled (Figure 1). The same trend is observed when considering several other measures of knowledge intensity, such as educational attainment, number of scientific publications, ratio of intangibles to assets, and share of workers employed in R&D activities and creative sectors (Florida, 2002). Proposed explanations for this structural shift include globalization, automation of routine jobs, and the steady increase in the burden of knowledge that demands an ever-increasing R&D effort to sustain a constant rate of productivity growth (Jones, 1995; Jones, 2009; Bloom et al., 2020).

There is a rich body of research that studies the impact of the increasing importance of knowledge-based activities on critical outcomes such as income inequality (Aghion et al., 2019; Gans and Leigh, 2019) and economic disparities across cities and regions (Moretti, 2012; Diamond, 2016; Gaubert et al., 2021). The widening economic gap between cities with strong knowledge-based economies and cities specialized in traditional industries is one of the most striking aspects of this transformation, with the former experiencing extraordinary growth in relative population and income compared to the latter (Glaeser and Gottlieb, 2009).

In this paper, we investigate whether the rise of the knowledge economy is causally linked to another major trend that has reshaped the economic geography of the United States in the last decades, namely, the marked increase of economic segregation within urban areas. Our preferred measure of economic segregation, the cross census tracts-within commuting zone Gini index, increased from 19.2 to 21.6 Gini points over the period 1990-2010, closely tracking the trend in overall income inequality that, according to the Census Bureau’s estimates, increased from 42.8 to 47.0 Gini points in the same period. To address this question empirically, we exploit the heterogeneous exposure of cities to the rise of the knowledge economy, and study the relationship between the local expansion in knowledge-based activities, measured through patenting growth, and changes in economic segregation from 1990 to 2010.\footnote{A rich literature has shown that economic segregation has a first-order impact on several policy-relevant local outcomes, including schooling choices (Katz et al., 2001; Baum-Snow and Lutz, 2011), health (Acevedo-Garcia et al., 2003; Alexander and Currie, 2017), and inter-generational mobility (Chetty et al., 2016; Fogli and Guerrieri, 2019).}

Theoretically, there are several reasons to presume the existence of this causal link. First, innovation and other creative jobs crucially depend on knowledge transmission, that an extensive literature has shown to be strongly localized (Glaeser et al., 1992; Jaffe et al., 1993; Carlino
and Kerr, 2015). This implies that an increase in the returns to accessing new ideas makes geographical clustering more attractive. Second, workers in the knowledge economy tend to be disproportionately sensitive to local amenities, such as quality of schooling, that reinforce the incentives for geographical segmentation (Baum-Snow and Hartley, 2020; Couture and Handbury, 2020). However, isolating the impact of an expansion in innovation-based activities on economic segregation is challenging because of potential reverse causation and the presence of unobservable factors, such as local financial or housing shocks, that jointly affect the urban environment and the ability of a geographical area to develop a knowledge-based economy.

To identify this causal relationship, we adopt an instrumental variable (IV) approach that exploits a newly assembled dataset of geo-referenced U.S. patents for the years 1975–2014. We propose that the production of local innovation is, at least in part, the result of a process of diffusion and recombination of ideas generated elsewhere in the economy, that propagate across geographical areas and technology classes through channels that are predetermined and persistent over time. To measure the strength of these channels, we leverage the observed network of citations in the early sample (1975-1994) and construct, for each city, a measure of local exposure to the emergence of new external ideas. We combine this measure with observed patenting in other locations and technology classes to obtain a prediction for a plausibly exogenous component of growth in local patenting between 1990 and 2010.²

The identifying assumption is that the predetermined linkages of knowledge diffusion are orthogonal, conditional on controls, to other local shocks that affect the evolution of economic segregation. We run an extensive set of robustness checks to gauge the credibility of this assumption and to address possible validity concerns. Particularly, among the other tests, we show that instrumented patenting growth is uncorrelated with previous trends in segregation and that our results are robust to controlling for the local industrial composition and local exposure to nation-wide technological trends.

Our main findings suggest that an expansion of local innovation activities, which we refer to as an innovation shock, leads to a significant increase in economic segregation. Our preferred estimates imply that a one standard deviation increase in patenting growth causes a 0.65 Gini point increase in income segregation, corresponding to 30.6% of a standard deviation in the 1990-2010 change in income segregation. We find consistent results when considering alternative measures of economic segregation by occupation and educational attainment. The analysis further reveals that the effect on income segregation cannot be fully explained by diverging income paths of initially segregated neighborhoods (inequality channel). A significant part of the effect is, in fact, explained by an increase in the geographical sorting of households along

²Our procedure extends the model proposed by Acemoglu et al. (2016) to a setting with multiple geographical locations. The strategy we propose is general and can be applied to other contexts in which channels of knowledge diffusion are measurable.
Notes: The blue line (left axis) is the contribution to U.S. GDP (value added) of computer and electronic products, electrical equipment, appliances and components, information, finance and insurance, professional and business services, educational services, health care and social assistance, arts, entertainment and recreation (data from the BEA). The dashed red line (right axis) is the number of patents per 1,000 people issued to U.S. inventors by the USPTO.

Why does an expansion in knowledge-based activities lead to more pronounced residential sorting within cities? In the second part of the paper, we use data from the National Establishments Time Series (NETS) to analyze changes in the location of jobs and consumption amenities in response to innovation shocks, and provide suggestive evidence of a mechanism that might be driving the estimated effects on economic segregation.

We first interpret local innovation shocks as an increase in the returns from localized knowledge spillovers, that reinforce the incentives of firms in knowledge-intensive sectors to co-locate. As proposed and extensively tested in the innovation literature, geographical proximity is a key determinant of the transmission of ideas, and learning externalities rapidly decline with distance (Jaffe et al., 1993, Carlino and Kerr, 2015, Catalini, 2018). For this reason, when new knowledge becomes available and gives rise to innovation opportunities, firms’ incentives to co-locate increase and, for the marginal firm, outweigh congestion costs. Consistently with this interpretation, we show that larger innovation shocks are associated to more pronounced clustering of knowledge-intensive jobs in neighborhoods characterized by a high initial density of workers in knowledge-intensive occupations (knowledge workers).

We then show that this clustering of employment coincides with changes in the residential choices of knowledge workers. In particular, in response to an expansion of local innovation...
activities, we find that knowledge workers relocate in the proximity of the census tracts that experience an inflow of knowledge-intensive jobs, presumably to reduce their commuting distance from these new employment opportunities. The effect declines sharply with space and even reverses for neighborhoods at more than 20 minutes of commuting distance. Additionally, we show that the rental price of housing and local consumption amenities (measured as the number of restaurants, food shops, and fitness centers) display an analogous spatial response, possibly amplifying the effect of employment clustering on economic segregation.

Related Literature

This study contributes to the literature on the determinants of urban segregation in the United States. Jargowsky (1996) documents a steady increase in economic segregation in U.S. metropolitan areas since 1970, and compares this trend with the slow decline in racial segregation. More recently, Reardon and Bischoff (2016) document that the trend in residential segregation that started in the 1980s continued, to a lesser extent, until very recently, and is correlated with the increase in income inequality. Baum-Snow and Pavan (2013) document a positive relationship between city size and the increase in the dispersion of earnings, and interpret this relation as evidence of a skill-biased change in agglomeration economies. Our paper is closely related to Diamond (2016) and Rossi-Hansberg et al. (2019), who study the growing geographical segmentation of individuals by skills across U.S. cities. Our focus is on the determinants of income, occupational, and educational sorting within cities.

A growing literature in urban economics investigates the relationship between recent changes in the internal structure of cities and the evolution of preferences, technology, and the income distribution. Couture et al. (2020) develop a model with non-homothetic preferences in which an increase in income inequality generates demand for high-quality amenities, leading to gentrification of downtown neighborhoods and amplifying the welfare disparities between high- and low-income households in cities. Couture and Handbury (2020) provide an empirical assessment of the drivers of downtown revivals since 2000 and find evolving preferences for downtown amenities among young college graduates to be the largest contributor. A complementary view is provided by Su (2020), who argues that the rising value of time for high-skilled workers increased their demand for central locations, contributing to downtown gentrification. Our paper provides evidence that the emergence of knowledge-based activities affects the geographical distribution of residents and jobs within cities. In particular, we find that changes in the location of consumption amenities play an important role in driving the increase in the sorting of residents that we document. This finding contributes to an expanding literature that studies the geography of consumption activities within cities. Recent contributions to this literature include Davis et al. (2019), who document the determinants of consumption segregation in
New York City using data from Yelp, Gorback (2020), who studies changes in the availability of consumption amenities in response to the entry of ridesharing platforms, and Almagro and Domínguez-Iino (2021), who use data from Amsterdam to analyze the endogenous supply of different types of amenities in a setting that allows for heterogeneous preferences among agents, and study the implications for residential sorting and welfare inequality. While the analysis in our paper is entirely reduced-form, in ongoing work (Berkes and Gaetani, 2021) we develop and estimate a structural model in the spirit of Ahlfeldt et al. (2015), that formalizes agents’ residential and commuting choices while accounting for the endogeneity in house prices and consumption amenities, and that we use to quantitatively disentangle the economic mechanisms underlying the link between innovation shocks and economic segregation.

This study also contributes to the expanding literature on the distributional effects of innovation. Aghion et al. (2019) exploit cross-state variation and find that changes in innovation intensity can explain the rise in top-income inequality in the United States. Jones and Kim (2018) formalize this link in the context of a Schumpeterian endogenous growth model. Florida and Mellander (2015) conduct a comprehensive study of urban segregation in U.S. metro areas and relate this increase to the expansion of jobs in the high-technology industry. In the present study, we provide causal evidence that supports their interpretation.

The rest of the paper is organized as follows. Section 2 presents the data sources and describes the measures of economic segregation used throughout the paper. Section 3 describes the main correlations and the IV strategy, and provides a discussion of the instrument and robustness results. Section 4 outlines and tests an economic mechanism that can explain the main results. Section 5 discusses avenues for future research and concludes.

2 Data and measurement

We combine data on innovation activities, captured by patenting, with social and economic indicators from the Census and the American Community Survey (ACS). For expositional purposes, we interpret commuting zones (CZs) as cities and census tracts (CTs) as neighborhoods (and use the terms interchangeably throughout the text). Commuting zones are defined with respect to actual commuting flows in the U.S. and, contrary to MSAs, constitute a complete partition of the country.\(^3\) Since our objective is to assess how innovation shocks affect the concentration of residents and workers within local labor markets, and those two aspects are intrinsically connected through commuting decisions, commuting zones are the natural unit of geographical aggregation for our analysis.

\(^3\)We use the definition of 1990 commuting zones provided by the Integrated Public Use Microdata Series (IPUMS, Ruggles et al., 2021). We only consider commuting zones in the contiguous United States.
2.1 Patents data

We use patenting at the city level to proxy for local knowledge intensity. Patents data are collected from the United States Patents and Trademark Office (USPTO). We parse the text of all the patents filed between 1975 and 2014 (and issued until 2015) and construct a new dataset that includes, for each grant, information on filing and issuing years, technology class, forward and backward citations as well as residence (city and state) of its inventors. Grants are then assigned to a city based on the location of their first inventor. From the publicly available documents, we identify a total of 5,030,264 patents, out of which 2,634,606 are located in the United States.\(^5\)

2.2 Measures of economic segregation

We construct measures of economic segregation that capture the concentration of residents by income, occupation, and education across neighborhoods within each city. Information on economic and demographic characteristics at the census tract level is taken from the National Historical Geographic Information System (NHGIS, Manson et al., 2020) that combines data from the decennial Census until 2000 and the American Community Survey (ACS) from 2005 onwards.\(^6\) Throughout the analysis, we keep census tract boundaries fixed at their 1990 definition, and use area-based crosswalks to assign values from other periods to the 1990 geography.

Income segregation is defined as the cross-neighborhood within-city Gini index, and is meant to capture the dispersion of income within a given city, once the variation of income within each neighborhood has been removed. Formally, letting \(\mathcal{N}_c = \{1, \ldots, N_c\}\) be the set of neighborhoods in city \(c\) ordered from the one with the lowest average household income to the highest, income segregation is defined as:

\[
IncSegr_c = 100 \times \left[1 - 2 \times \sum_{n=1}^{N_c} \left(\frac{h_n}{H_c} \sum_{n'=1}^{n} \frac{x_{n'}}{X_c}\right)\right],
\]

where \(h_n\) and \(H_c\) denote the number of households, and \(x_n\) and \(X_c\) total income of neighborhood \(n\) and city \(c\), respectively. It is easy to show that, in the extreme case in which the average

---

\(^4\)Our analysis is based on the International Patent Classification (IPC). We assign to each patent a technology class using a frequency-based crosswalk from its main USPTO class to the corresponding class in the IPC.

\(^5\)The reported residence of the first inventor of the remaining patents is located outside of the United States.

\(^6\)We use the 2008-2012 ACS averages to generate economic and demographic variables for the 2010 observation. Income data at the CT level divide households into 15 brackets, with lower bounds \(0\), \(10,000\), \(15,000\), \(20,000\), \(25,000\), \(30,000\), \(35,000\), \(40,000\), \(45,000\), \(50,000\), \(60,000\), \(75,000\), \(100,000\), \(125,000\), and \(150,000\). Computing inequality measures from binned data is made problematic by the fact that the top bin is unbounded, with an average that potentially varies substantially across census tracts. The literature has approached this issue in different ways, each with its own advantages and limitations. Appendix B.1 discusses them and provides a detailed description of the procedure we use to approximate the income distribution.
income of each neighborhood is the same, \( \text{IncSegr}_c \) is equal to zero. On the contrary, if households are perfectly sorted across neighborhoods, income segregation is equal to the standard within-city Gini index.\(^7\)

We measure segregation by occupation and education as the index of dissimilarity,\(^8\) that is meant to capture the asymmetry across neighborhoods in the prevalence of a group of individuals. In our case, we define these groups as residents employed in knowledge-intensive occupations (for occupational segregation), and residents with at least a 4-year college degree (for educational segregation). Formally, we define occupational segregation as:

\[
\text{OccSegr}_c = 100 \times \sum_{n=1}^{N_c} \frac{1}{2} \left| \frac{r^1_n}{R^1_c} - \frac{r^0_n}{R^0_c} \right|,
\]

where \( r^1_n \) and \( R^1_c \) denote the number of residents employed in knowledge-intensive occupations (knowledge workers) in neighborhood \( n \) and city \( c \), respectively, and \( r^0_n \) and \( R^0_c \) denote the corresponding number of residents employed in residual occupations. Appendix B.2 provides details on the definition of knowledge-intensive occupations and the construction of the numbers of workers by type at the census tract level. Analogously, we define educational segregation as:

\[
\text{EduSegr}_c = 100 \times \sum_{n=1}^{N_c} \frac{1}{2} \left| \frac{r^H_n}{R^H_c} - \frac{r^L_n}{R^L_c} \right|,
\]

where the superscript \( H \) denotes educational attainment of at least a 4-year college degree, and \( L \) any educational attainment below a 4-year college degree.

Appendix Figure A.1 displays histograms of the distribution of the three measures of economic segregation in 1990 (left panels), 2010 (middle panels), and the change between 1990 and 2010 (right panels). The evolution of these measures exhibits considerable variation across cities. The change in income segregation has an interquartile range of 2.5 Gini points, approximately 13% of the nationwide average in 1990 (19.2). Changes in occupational and educational segregation display a comparable degree of variation, with interquartile ranges of 3.8 and 3.2, respectively, equal to 17.6% and 10.1% of their nationwide averages in 1990 (21.5 and 32.0, respectively). Appendix Table A.2 reports the corresponding numbers for the 15 largest commuting zones. Between 1990 and 2010, some of the major cities (including San Francisco, Boston, and Seattle) experienced increases in income segregation of more than 5 Gini points, while other major urban areas (including Pittsburgh and Houston) recorded considerably smaller increases of less than 2 Gini points.

\(^7\)In the implementation of Equation (1), we use a piecewise linear, instead of a step function, to approximate the Lorenz curve. This guarantees that \( \text{IncSegr}_c \) is always between zero and one.

\(^8\)See White (1983) for a discussion of the pros and cons of the index of dissimilarity as a measure of segregation.
2.3 Other data sources

We complement our analysis with data on the geographical distribution of workers by occupation (at the place of employment), the availability of local consumption amenities, cross-census tract bilateral commuting times, and local rental price of housing.

Data on employment and local amenities is compiled from the National Establishment Time Series (NETS). The NETS provides data on employment, geographical location and industry for close to the universe of U.S. establishments over the period 1990-2015.\(^9\)

Bilateral commuting times across census tracts are computed using the Open Source Routing Machine (OSRM).\(^10\) This routing engine allows us to compute travel time by car for each pair of neighborhoods within each city, for a total of almost 19.4 million pairs.

Finally, the rental price of housing is computed from the NHGIS as average rent per room.\(^11\) Appendix Table A.1 and Appendix B provide summary statistics and further details on the construction of the main variables.

2.4 Data timeline

For most of the analysis, we study changes in local outcomes over a 20-year period (specifically, between 1990 and 2010). To avoid capturing transitory shocks to innovation, we measure patenting activity for each time period (1990 and 2010) as ten-year totals around the focal year (1985-1994 and 2005-2014, respectively).

Patents data cover a 40-year period that we divide into two 20-year samples. The early sample (1975-1994) is used in the IV analysis to infer knowledge links across geographical areas and technology classes and to measure innovation for the 1990 observation. The late sample (1995-2014) is itself divided into two time periods. The first decade (1995-2004) is used in conjunction with the knowledge links to calculate the local exogenous shocks to innovation. The second decade (2005-2014) is used to measure innovation for the 2010 observation. The time structure of the data is illustrated in Figure 2.

\(^9\)More details on the NETS data and the procedure to assign each establishment to a census tract can be found in Appendix B.2.
\(^10\)http://project-osrm.org/
\(^11\)Specifically, average rent per room is computed as aggregate gross rent for renter-occupied housing units paying cash rent divided by the aggregate number of renter occupied rooms.
The impact of innovation activities on economic segregation

Does an expansion in knowledge-intensive activities induce an increase in economic segregation within cities? To answer this question, we first identify a causal link between those phenomena and investigate its features. We then provide evidence for a plausible mechanism underlying this relationship.

The empirical model relates changes in economic segregation and patenting growth at the city level between 1990 and 2010:

$$\Delta Y_c = \alpha + \beta \Delta \log (1 + \text{Patents}_c) + \gamma Z_{c,1990} + \epsilon_c, \quad (4)$$

where $Z_{c,1990}$ represents a set of city-level controls at their 1990 values. Throughout the paper, we report the estimates of the coefficient of interest, $\beta$, with the outcome variable, $\Delta Y_c$, representing the change in each of the three measures of economic segregation defined in Section 2.2. In the regressions, we weight cities by the number of households in 1990.\footnote{In Appendix Table A.5 we report OLS and two-stage least squares (2SLS) results when regressions are unweighted but only include cities with 1990 number of households above 60,000. The resulting sample, that includes 259 cities, accounts for 88.8\% of total U.S. households in 1990. Results are strongly consistent.} To avoid dropping observations with zero patents either in 1990 or 2010, we adopt the convention of taking the logarithm of one plus total patents.\footnote{In Table A.6, we show that results are robust to using an inverse hyperbolic sine (arcsinh) transformation. Since zeros in patent counts are concentrated in scarcely populated areas and observations are weighted by population, we obtain virtually identical results when adopting other strategies used in the literature (e.g.,}
Figure 3: **Patenting growth and economic segregation: unconditional correlations**

*Notes:* Scatter plots of the unconditional correlations between growth in patenting and change in income segregation (left panel), occupational segregation (middle panel), and educational segregation (right panel) between 1990 and 2010. Circles and regression lines are weighted by the total number of households in 1990.

### 3.1 Correlations and OLS

The left panel of Figure 3 shows a scatter plot of the unconditional correlation between changes in income segregation and the growth rate of total patents between 1990 and 2010 (the size of each circle is proportional to the total number of households in 1990). The $R^2$ of the underlying weighted regression is 0.08 and the estimated coefficient is 1.21, implying that a 10 percentage point increase in patenting growth is associated with a 0.121 Gini point increase in the change in income segregation. The center and right panels of Figure 3 show analogous positive correlations for the measures of occupational and educational segregation.

Table 1 reports the regression results, for each of the three measures of segregation, once we control for potential confounding factors, including the share of individuals with at least a 4-year college degree, log number of census tracts, log number of households, and log average income in 1990. These controls account for the fact that the increase in innovation activities and income segregation is likely to be correlated with the overall level of human capital, city size, economic conditions, and for the fact that cities with a higher number of census tracts may have mechanically higher measures of segregation. Local industry composition could also be a source of bias if aggregate shocks at the industry level (notably, trade shocks) had an impact on both the expansion of local knowledge-intensive activities and other variables affecting the urban environment. To account for this possibility, we control for trade shocks using the measure of exposure to import from China developed by Autor et al. (2013).

The inclusion of the full set of controls significantly dampens the estimated effect of patenting growth on all the three measures of segregation. While the coefficient remains significant for occupational and educational segregation (Panels B and C), it loses statistical significance including dummy variables for zeros, or taking growth rates through midpoint method). Since measures of patenting activity are based on ten-year totals, only 16 commuting zones have no patenting either in 1990 or in 2010. The combined number of households of those cities is about 0.04% of total U.S. households in 1990.
### Table 1: Patenting growth and economic segregation: OLS results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: ( \Delta \text{IncSegr} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patenting growth</td>
<td>1.21***</td>
<td>0.27</td>
<td>0.40</td>
<td>0.11</td>
<td>0.16</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(0.36)</td>
<td>(0.31)</td>
<td>(0.32)</td>
<td>(0.30)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>16.60***</td>
<td>7.23*</td>
<td>8.27**</td>
<td>-0.22</td>
<td>-0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.87)</td>
<td>(4.26)</td>
<td>(3.96)</td>
<td>(4.63)</td>
<td>(4.55)</td>
<td></td>
</tr>
<tr>
<td>Log CTs</td>
<td>0.55***</td>
<td>-2.35**</td>
<td>-2.45**</td>
<td>-2.43**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(1.09)</td>
<td>(1.17)</td>
<td>(1.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log households</td>
<td>2.80**</td>
<td>2.54**</td>
<td>2.53*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.11)</td>
<td>(1.25)</td>
<td>(1.27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log average income</td>
<td>4.55***</td>
<td>4.52***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.34)</td>
<td>(1.34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Import exposure</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R(^2)</strong></td>
<td>0.08</td>
<td>0.25</td>
<td>0.33</td>
<td>0.38</td>
<td>0.40</td>
<td>0.41</td>
</tr>
</tbody>
</table>

|                  | (1)     | (2)     | (3)     | (4)     | (5)     | (6)     |
| **Panel B: \( \Delta \text{OccSegr} \)**                   |         |         |         |         |         |         |
| Patenting growth | 1.98*** | 1.00*   | 1.20**  | 1.16**  | 1.21*** |
|                  | (0.56)  | (0.56)  | (0.51)  | (0.46)  | (0.44)  |
| Share of college graduates | 17.47*** | 2.19    | 2.34    | -5.98   | -6.08   |
|                  | (4.15)  | (3.59)  | (3.57)  | (4.96)  | (4.99)  |
| Log CTs          | 0.90*** | 0.48    | 0.38    | 0.38    |
|                  | (0.12)  | (0.74)  | (0.78)  | (0.77)  |
| Log households   | 0.41    | 0.15    | 0.16    |
|                  | (0.71)  | (0.70)  | (0.70)  |
| Log average income| 4.47*   | 4.48*   |
|                  | (2.27)  | (2.28)  |
| Import exposure  | -0.03   |
|                  | (0.12)  |
| **R\(^2\)**     | 0.13    | 0.24    | 0.37    | 0.37    | 0.38    | 0.38    |

|                  | (1)     | (2)     | (3)     | (4)     | (5)     | (6)     |
| **Panel C: \( \Delta \text{EduSegr} \)**                   |         |         |         |         |         |         |
| Patenting growth | 1.94*** | 1.16**  | 1.27**  | 1.15**  | 1.24*** |
|                  | (0.60)  | (0.56)  | (0.55)  | (0.48)  | (0.43)  |
| Share of college graduates | 13.97*** | 5.36    | 5.80    | -9.96   | -9.32   |
|                  | (3.83)  | (3.91)  | (3.77)  | (6.03)  | (5.87)  |
| Log CTs          | 0.51**  | -0.72   | -0.90   | -0.85   |
|                  | (0.22)  | (1.20)  | (1.19)  | (1.19)  |
| Log households   | 1.18    | 0.70    | 0.68    |
|                  | (1.19)  | (1.11)  | (1.12)  |
| Log average income| 8.45*** | 8.36*** |
|                  | (2.60)  | (2.53)  |
| Import exposure  | 0.21    |
|                  | (0.14)  |
| **R\(^2\)**     | 0.12    | 0.18    | 0.22    | 0.22    | 0.28    | 0.28    |

|                  |         |         |         |         |         |         |
| **# Obs.**       | 722     | 722     | 722     | 722     | 722     | 722     |

**Notes:** Regressions are weighted by total number of households in 1990. Controls are at 1990 values, with the exception of Import exposure (defined as in Autor et al., 2013). Standard errors clustered at the state level in parenthesis. *** \(p < 0.01\); ** \(p < 0.05\); * \(p < 0.1\).
in the case of income segregation (Panel A). The loss of significance is due to the inclusion of the variables that control for size or human capital, possibly reflecting the fact that larger cities, with a higher share of college-educated individuals, have experienced in recent decades a more pronounced increase in income inequality, as documented by Baum-Snow and Pavan (2013). Since income segregation is intrinsically connected to inequality, as we explore in depth in Section 3.5, and patenting has grown more in larger cities, it is not surprising that the OLS estimate of $\beta$ decreases once we control for these city characteristics.

### 3.2 Instrumenting for patenting activity

We now move to isolate the causal relationship between growth in innovation activities and economic segregation. To this end, we need to identify variation in patenting that, conditional on controls, is orthogonal to unobserved factors that might affect at the same time the expansion of a knowledge-based economy and changes in segregation in cities. The range of such possible factors is large and the direction of the bias is ex-ante ambiguous. Examples of such factors include short-run phenomena such as housing and financial shocks, or long-run trends such as technological obsolescence of local industries. Inverse causality is also a possible concern, with segregation potentially being the cause, rather than the consequence, of the emergence of the knowledge economy in U.S. local labor markets.

In this section, we propose an instrument for the local growth in innovative activities. The intuition behind the instrument is that local patenting is, at least in part, the result of a process of recombination of ideas generated elsewhere in the economy. These ideas propagate across geographical areas and technology classes through channels of diffusion that are predetermined and persistent over time. This process is conceptually similar to an input-output model for the production of ideas, in which existing patents are perfectly substitutable building blocks for future innovation.\[^{14}\] We use the observed network of patent citations in the early sample (1975-1994) to measure the strength of the channels of knowledge diffusion underlying this process. By percolating actual patenting in the period 1995-2004 through these diffusion channels, we obtain a prediction for local patenting in 2005-2014.

To illustrate the procedure we use to predict patenting in each location, consider the fictitious example displayed in Figure 4. The economy is composed of 3 city-class pairs: Detroit-Vehicles, San Francisco-Computers, and Chicago-Metallurgy. The left panel depicts the network of citations observed in the early sample. Detroit-Vehicles produces one patent that cites 3 grants (out of 100) from San Francisco-Computer and 3 more (out of 5) from Chicago-Metallurgy. Under the assumption that external ideas are perfectly substitutable inputs for

\[^{14}\text{The main departure from a traditional input-output model of production is that in our case ideas are non-rival inputs. As a result, the sum of all the inputs that appear in the production of new patents is, in general, larger than the overall amount of available inputs.}\]
Notes: Illustrative example of the diffusion and recombination process underlying the construction of the instrument (for clarity, the example abstracts from diffusion lags). Left panel: In the early sample, the destination city-class Detroit-Vehicles produces one patent, that cites 3 patents (out of 100) from the origin city-class San Francisco-Computers and 3 patents (out of 5) from Chicago-Metallurgy. This yields coefficients of diffusion equal to $\frac{0.5}{100}$ and $\frac{0.5}{5}$, respectively. Right panel: In the late sample, Chicago-Metallurgy produces 5 additional patents. Our procedure predicts an additional half patent in Detroit-Vehicles.

local innovation, these observed citation links suggest that we need 100 new patents in San Francisco-Computer, or 5 new patents in Chicago-Metallurgy (the origin city-class pairs), to generate half of a new patent in Detroit-Vehicles (the destination city-class pair). In the formalization of the instrument, we will refer to the quantity of patents in the destination city-class induced by an additional patent in the origin city-class as the coefficient of diffusion. In this case, the coefficients of diffusion are equal to $\frac{0.5}{100}$ and $\frac{0.5}{5}$ for San Francisco-Computer and Chicago-Metallurgy, respectively, where 0.5 is the share of citations allocated by Detroit-Vehicles to each origin city-class. The right panel illustrates how we use these coefficients to predict patenting in the late sample: as Chicago-Metallurgy produces 5 additional patents, we predict an additional half grant ($\frac{0.5}{5} \times 5$) in Detroit-Vehicles.

This example illustrates the basic intuition behind the instrument. In our implementation, we generalize this procedure by taking into account the time lag between the cited and citing patent. In the following sub-section, we formalize this intuition and describe our procedure in detail.

### 3.2.1 Construction of the instrument

As illustrated in the example above, we build the instrument in two steps. In the first step, we use the observed citation patterns in the early sample (1975-1994) to identify and quantify the channels of idea diffusion across cities and technology classes. For every patent produced in city $c$ of class $\nu$ (the destination city-class pair), we calculate the share of citations given to
patents from city \( b \) of class \( \mu \) (the origin city-class pair) filed \( \tau \) years before. We interpret the sum of all these shares, normalized by the total number of patents in \( (b, \mu) \), as a measure of the strength of knowledge flows from the origin city-class to the destination city-class. We call this measure the coefficient of diffusion and denote it by \( d_{(b,\mu)\rightarrow(c,\nu)}^\tau \). This term captures the number of patents induced in \( (c, \nu) \) by one patent filed \( \tau \) years before in \( (b, \mu) \).\(^{15}\)

To compute the coefficients of diffusion, we consider each patent filed between 1985 and 1994 as a potential destination patent, and those filed between 1 and 10 years before as potential origin patents.\(^{16}\) Mathematically,

\[
d_{(b,\mu)\rightarrow(c,\nu)}^\tau = \begin{cases} \sum_{p \in P(c,\nu)} \frac{ShareCit_{(b,\mu)\rightarrow p}^{\tau}}{TotPat_{b,\mu}^{\tau}} & \text{for } b \neq c, \\ 0 & \text{for } b = c \end{cases} \quad \text{for } \tau \in \{1, \ldots, 10\}, \quad (5)
\]

where \( P(c,\nu) \) is the set of patents in the destination city-class \( (c, \nu) \) filed between 1985 and 1994. Equation (5) defines the coefficient of diffusion as the ratio of two quantities. The numerator is the sum of the share of citations given by each destination patent \( p \in P(c,\nu) \) to origin patents in \( (b, \mu) \) filed \( \tau \) years before \( (ShareCit_{(b,\mu)\rightarrow p}^{\tau}) \). The denominator \( (TotPat_{b,\mu}^{\tau}) \) is the total number of potential origin patents in \( (b, \mu) \) filed between 1985\( - \tau \) and 1994\( - \tau \). To reduce endogeneity concerns, we set the coefficient for links that start and end in the same city to zero.

In the second step, we use the coefficients of diffusion in combination with observed patenting in the intermediate period (1995-2004) to predict local patenting between 2005 and 2014. We do this by computing the number of patents that we expect to observe in each city given the input-output model outlined above. Formally, to predict patenting in city \( c \) in 2005, we multiply the observed patenting activity of each city-class pairs in 2005\( - \tau \) by the relevant coefficients for \( \tau = 1, \ldots, 10 \), and take the sum of the resulting numbers:

\[
\hat{Pat}_{c,2005} = \kappa_{2005} \sum_{\tau=1}^{10} \sum_{b,\mu,\nu} d_{(b,\mu)\rightarrow(c,\nu)}^\tau \Pat_{b,\mu,2005\rightarrow \tau}. \quad (6)
\]

The term \( \kappa_{2005} \) is a rescaling factor that ensures that the total number of patents we estimate nationwide is the same as the one we observe in the data. The same strategy is used to predict patenting in 2006, with the only exception that the coefficients at \( \tau = 1 \) are applied to the predicted patents in 2005, instead of the actual ones.\(^{17}\) We do this to avoid endogeneity concerns.

\(^{15}\)Note that we refer to a cited patent as an idea origin, and to a citing patent as an idea destination. The direction of the arrow denotes knowledge flows from the origin (cited) to the destination (citing) patent.

\(^{16}\)This implies that all patents filed between 1975 and 1993 are potential origin patents. Note that, in general, destination patents can also be origin patents.

\(^{17}\)The role of \( \kappa_{2005} \) is now evident: It prevents the national predicted number of patents in the later years to
that might arise when using contemporaneous patenting. This process is repeated sequentially to predict patenting for all years up to 2014. Appendix Table A.3 illustrates the timing of this sequential procedure.

The instrument for local patenting growth is defined as the log difference between predicted patents in the late sample (2005-2014) and actual patents in the early sample (1985-1994).\textsuperscript{18} In the remainder of the paper, we refer to this measure as predicted patenting growth or, interchangeably, as \textit{innovation shock}.

### 3.2.2 Conditions for validity

To be able to identify the causal effect of innovative activities on economic segregation, the usual conditions for instrument validity must hold.\textsuperscript{19} First, the instrument must have predictive power on actual patenting growth between 1990 and 2010. We extensively discuss this point in the next sub-section, where we show that the network of diffusion inferred in the early sample is in fact persistent and is successful at predicting innovation in the late sample. Second, the instrument must isolate variation in patenting growth that, conditional on controls, is uncorrelated with unobservable factors that affect the trajectory of local segregation. It is important to stress that, as common in IV settings used in the economic geography literature (e.g., Baum-Snow, 2007, Duranton and Turner, 2012, and Agrawal et al., 2017), the orthogonality requirement must hold \textit{conditionally} on the set of local characteristics that might affect the formation of knowledge linkages in the early sample and might be correlated with the evolution of local segregation since 1990. In particular, it is plausible that cities with a higher local density of human capital may be endowed with more valuable knowledge linkages (possibly because of the presence of universities or large innovative firms) and, at the same time, experience systematically different trends in economic segregation since 1990. For this reason, we argue that our instrument is valid conditional on the initial local share of college graduates. We show that, conditional on this control, the instrument is uncorrelated with local pre-trends in segregation and the estimates pass the test of coefficient stability proposed by Oster (2019). Moreover, our results are robust to a wide range of alternative specifications of the empirical model and variations of the instrument.

A further important concern for the validity of the instrument is that the channels of diffusion measured through Equation (5) reflect a demand-pull from the destination city-class

\textsuperscript{18} Recall that patenting in the 1990 (2010) observation is defined as the sum of actual or predicted patents in the years 1985-1994 (2005-2014).

\textsuperscript{19} Note that our instrument falls into the family of shift-share instruments, as it combines a predetermined network of knowledge links (shares) and observed patenting activity across cities (shifts). The conditions for the validity of this class of instruments are discussed by Goldsmith-Pinkham et al. (2020) and Borusyak et al. (2018).
rather than a supply-push from the origin city-class. In other words, it is possible that cited inventions are produced because of demand factors in the citing city. In this case, patenting in the origin city-class would be itself an outcome of local shocks in the destination city-class, leading to a violation of the orthogonality condition. To rule out this possibility, in Appendix C, we perform a test, in the spirit of Acemoglu et al. (2016), where we exploit the asymmetric nature of the citation network and we predict patenting in 1995-2004 using, alternatively, “supply” links from 1985-1994 (using backward citations) and “demand” links from 2005-2014 (using forward citations). We show that, conditional on the supply links, demand links do not have predictive power on actual patenting.

Taken together, these results suggest that our estimates are unlikely to be affected by unobservable factors and reflect the causal impact of patenting growth on the evolution of local segregation.

3.2.3 First-stage results

The first condition for the instrument to be valid is that predicted patenting growth, as obtained through the procedure outlined above, is correlated with actual patenting growth. This condition requires that the network of knowledge diffusion inferred from the citation patterns is stable over time, which we can test directly by comparing the network in the early sample with its counterpart in the late sample. In Appendix D, we compute the distance between the two observed networks, and contrast it with the distance between pairs of analogous networks in which individual patents have been randomly reallocated to cities in such a way as to maintain the total number of grants for each city-year unchanged. The results show that the two observed networks are significantly closer to each other than the simulated ones, suggesting that, to some degree, the network retains its structure over time.

Figure 5 shows a scatter plot of the first stage relationship between predicted and actual growth rate of patenting. We plot the residuals of a regression of patenting growth on the full set of controls. The two variables are strongly but not perfectly correlated. The residual $R^2$ is 0.18, while the coefficient of the regression is 0.59. The F-statistic is 34.7 which rules out weak instrument concerns. Appendix Figure A.2 visually compares actual and predicted patenting growth at the city level on a map of the United States.

3.2.4 Pre-trend analysis

The second condition for the instrument to be valid is that cities that experienced different innovation shocks, as captured by the instrument, would have not followed systematically different trajectories in economic segregation in the absence of the shocks. We provide evidence in support of this condition by showing that predicted patenting growth is uncorrelated with
Figure 5: **First stage scatter-plot**

\[
\begin{align*}
\hat{y} &= 0.594 (0.047) \\
R^2 &= 0.181
\end{align*}
\]

Notes: Correlation between actual and predicted growth in patenting between 1990 and 2010, after residualizing with respect to the full set of controls (as in column 6 of Table 1). Circles and regression line are weighted by the total number of households in 1990.

pre-existing local trends in segregation. In Section 3.4, we will run additional robustness checks to address remaining endogeneity concerns.

The right panels of Figure 6 show reduced-form relationships between predicted patenting growth (1990-2010) and past changes in the three measures of economic segregation (1980-1990) once we partial out the full set of controls. For all the three measures, the slope of the regression line is statistically indistinguishable from zero. This suggests that innovation shocks in 1990-2010 are not correlated with trajectories of economic segregation in the previous decade.

Appendix Table A.4 reports estimates of the corresponding OLS regressions, where we progressively introduce the set of controls. The estimated coefficients of predicted patenting growth are statistically indistinguishable from zero at the 5% level for all the specifications. Note that the years we select to calculate past changes in segregation are dictated by data availability. In the 1980 Census, data at the census tract level are not available for the entirety of the United States, but only for the most densely populated areas. For this reason, the number of observations in the pre-trend analysis is lower. Our main results are robust to restricting the sample to cities for which 1980 data are available.

### 3.3 IV results

We now use our instrument to explore the causal effects of innovation shocks on economic segregation within cities. The left panels of Figure 6 display the reduced form relationships
Figure 6: Predicted patenting growth and economic segregation

Notes: Reduced-form relationship between predicted patenting growth and change in income segregation (top panels), occupational segregation (middle panels), and educational segregation (bottom panels) between 1980 and 1990 (left panels) and 1990 and 2010 (right panels), after residualizing with respect to the full set of controls (as in column 6 of Table 1). Circles and regression line are weighted by the total number of households in 1990.

between predicted patenting growth and changes in the three measures of economic segregation between 1990 and 2010, after partialling out the full set of controls. The plots show strong positive correlations for all the three measures, suggesting that cities with more favorable innovation shocks experienced significantly larger changes in economic segregation.

Table 2 reports the two-stage least squares (2SLS) estimates of the empirical model in Equation (4). As in the OLS regressions, we weight observations by total number of households in 1990. In this case, the coefficient of patenting growth (Panel A) remains positive and statistically significant when the full set of controls is included (column 6, our preferred specification).

The estimated coefficient in column 6 of Panel A implies that increasing 1990-2010 patenting growth by 10 percentage points leads to a 0.13 Gini point increase in income segregation. Since the standard deviation of patenting growth is 49.9% and the standard deviation in the change in income segregation is 2.12 Gini points, the estimated effect is economically meaningful. A one standard deviation increase in patenting growth causes a 0.65 Gini point increase in
income segregation, equal to 30.6% of a standard deviation in the 1990-2010 change in income segregation. The effect is even more pronounced for the measures of occupational (53.9% of a standard deviation) and educational (39.8% of a standard deviation) segregation.

The 2SLS estimates are larger than the ones in the OLS regressions. This fact suggests that unobservable factors affecting at the same time innovation and segregation tend to operate on the two variables in opposite directions. For example, negative shocks to the local financial sector may generate turmoil on the urban structure by tightening residential mortgages and, at the same time, limiting credit to innovative startups. Similarly, obsolescence of the local industry may result in lower patenting and an increase in unemployment or poverty, which increases economic segregation.

The inclusion of the control for human capital (column 2) attenuates the magnitude of the estimates, suggesting that, as discussed in Section 3.2.2, this variable is indeed informative of both the citation network in the early sample and factors that control the evolution of segregation between 1990 and 2010. However, the inclusion of the other controls does not meaningfully affect the size of the coefficients, indicating that our estimates are unlikely to be affected by unobservable confounding factors.

We can formally verify this by performing, on the reduced-form regression, the coefficient stability test developed by Oster (2019). Intuitively, this test proposes that when the inclusion of controls generates significant drops in the estimated coefficient of interest and small improvements in the $R^2$, the omitted variable bias on the estimates is potentially large. Conversely, when the inclusion of controls induces a meaningful increase in the $R^2$ without a significant reduction in the estimated coefficient, the omitted variable bias is likely to be small. We implement this test for the three measures of segregation using the specification of column 2 as our baseline. Setting the threshold to its recommended value of 1.3 times the $R^2$ of the corresponding regression with the full set of controls, the test yields values of 2.23, 198.17, and 6.95 for the measures of income, occupational, and educational segregation, respectively. These statistics capture the extent to which selection on observables needs to translate into selection on unobservables in order to generate a treatment effect of zero. All values are well above the critical value of one, indicating that selection on unobservables would need to be larger than selection on observables to generate a null treatment effect.\footnote{The corresponding test using column 1 as the baseline still delivers values above the critical value of 1 (namely, 1.11, 1.87, and 1.61 for the measures of income, occupational, and educational segregation, respectively).}

### 3.4 Robustness checks

In Table 3 we run an extensive set of robustness checks to address possible endogeneity concerns related to our IV approach, focusing on the measure of income segregation. First, in column
Table 2: Patenting growth and economic segregation: IV results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: ΔIncSegr</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patenting growth</td>
<td>3.22***</td>
<td>1.92*</td>
<td>1.77**</td>
<td>1.25**</td>
<td>1.34**</td>
<td>1.30**</td>
</tr>
<tr>
<td>(0.59)</td>
<td>(1.00)</td>
<td>(0.90)</td>
<td>(0.60)</td>
<td>(0.59)</td>
<td>(0.61)</td>
<td></td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>10.03*</td>
<td>0.93</td>
<td>3.14</td>
<td>-6.18</td>
<td>-5.91</td>
<td></td>
</tr>
<tr>
<td>(5.56)</td>
<td>(6.28)</td>
<td>(4.82)</td>
<td>(5.67)</td>
<td>(5.72)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log CTs</td>
<td>0.60***</td>
<td>-1.57</td>
<td>-1.65</td>
<td>-1.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.19)</td>
<td>(1.02)</td>
<td>(1.08)</td>
<td>(1.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log households</td>
<td>2.09**</td>
<td>1.79</td>
<td>1.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.02)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log average income</td>
<td>4.90***</td>
<td>4.87***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.63)</td>
<td>(1.62)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Import exposure</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.10)</td>
</tr>
<tr>
<td><strong>Panel B: ΔOccSegr</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patenting growth</td>
<td>4.18***</td>
<td>2.89***</td>
<td>2.65***</td>
<td>2.81***</td>
<td>2.90***</td>
<td>2.96***</td>
</tr>
<tr>
<td>(0.53)</td>
<td>(0.77)</td>
<td>(0.66)</td>
<td>(0.65)</td>
<td>(0.62)</td>
<td>(0.61)</td>
<td></td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>9.93**</td>
<td>-4.47</td>
<td>-5.13</td>
<td>-14.57***</td>
<td>-15.04***</td>
<td></td>
</tr>
<tr>
<td>(4.61)</td>
<td>(3.72)</td>
<td>(3.87)</td>
<td>(4.74)</td>
<td>(4.98)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log CTs</td>
<td>0.95***</td>
<td>1.61*</td>
<td>1.53</td>
<td>1.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.12)</td>
<td>(0.90)</td>
<td>(0.95)</td>
<td>(0.95)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log households</td>
<td>-0.63</td>
<td>-0.94</td>
<td>-0.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.86)</td>
<td>(0.90)</td>
<td>(0.90)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log average income</td>
<td>4.96**</td>
<td>5.00**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.23)</td>
<td>(2.26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Import exposure</td>
<td>-0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: ΔEduSegr</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patenting growth</td>
<td>3.62***</td>
<td>2.59***</td>
<td>2.37***</td>
<td>2.24***</td>
<td>2.41***</td>
<td>2.27***</td>
</tr>
<tr>
<td>(0.59)</td>
<td>(0.90)</td>
<td>(0.83)</td>
<td>(0.78)</td>
<td>(0.70)</td>
<td>(0.72)</td>
<td></td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>8.62*</td>
<td>0.35</td>
<td>0.89</td>
<td>-15.85***</td>
<td>-14.71***</td>
<td></td>
</tr>
<tr>
<td>(4.83)</td>
<td>(4.97)</td>
<td>(5.06)</td>
<td>(5.48)</td>
<td>(5.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log CTs</td>
<td>0.55**</td>
<td>0.03</td>
<td>-0.11</td>
<td>-0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.22)</td>
<td>(1.35)</td>
<td>(1.38)</td>
<td>(1.40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log households</td>
<td>0.50</td>
<td>-0.05</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.31)</td>
<td>(1.29)</td>
<td>(1.32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log average income</td>
<td>8.79***</td>
<td>8.67***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.47)</td>
<td>(2.42)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Import exposure</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Obs.</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>722</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First stage estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted patenting growth</td>
<td>0.71***</td>
<td>0.64***</td>
<td>0.65***</td>
<td>0.59***</td>
<td>0.59***</td>
<td>0.59***</td>
</tr>
<tr>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>F-stat</td>
<td>54.9</td>
<td>37.5</td>
<td>35.4</td>
<td>36.1</td>
<td>34.0</td>
<td>34.7</td>
</tr>
<tr>
<td>R²</td>
<td>0.39</td>
<td>0.39</td>
<td>0.40</td>
<td>0.41</td>
<td>0.41</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Notes: Regressions are weighted by total number of households in 1990. Controls are at 1990 values, with the exception of Import exposure (defined as in Autor et al., 2013). Standard errors clustered at the state level in parenthesis. ***p < 0.01; **p < 0.05; *p < 0.1.
we control for past trends in innovation (patenting growth between 1980 and 1990), that are weakly negatively correlated with the change in segregation, and induce a slight increase in the estimated coefficient of current patenting growth. Second, we address the critical concern that geographical areas linked in the knowledge network have characteristics, such as a similar industry structure, geographical proximity, common regulation, or exposure to other shocks, that make it hard to disentangle the genuine effect of knowledge shocks from the effect of other factors that affect, at the same time, innovation in the origin city and segregation in the destination city. To control for the effect of nationwide industry or technology-specific shocks, we include a Bartik-like variable in the set of controls. Namely, for each city $c$ we define a vector $\Lambda_c^{1990} = \{\lambda_{c,1}^{1990}, \ldots, \lambda_{c,S}^{1990}\}$, where $\lambda_{c,\nu}^{1990}$ denotes the share of patents produced in $c$ that belong to technology class $\nu$. We then compute the 1990-2010 growth rate $g_{-c,\nu}$ of the number of grants in $\nu$, considering only patents produced outside $c$. Our control variable is then computed as:

$$\hat{g}_c = \sum_{\nu=1}^{S} \lambda_{c,\nu}^{1990} \cdot g_{-c,\nu}. \tag{7}$$

This prediction replicates the idea behind a Bartik instrument, with the distribution of patents across technology classes used in place of the distribution of employment across industries. Column 3 shows the 2SLS regression once the Bartik-like variable is included in the set of controls. The coefficient on patenting growth remains positive and only marginally smaller in magnitude. A related concern is that a few technological areas (e.g., transportation, IT, etc...) receive a disproportionate share of knowledge flows from other fields, and their local prevalence is correlated with other factors driving changes in segregation. To address this concern, in column 4 we show that results are robust when controlling directly for the local share of patents in each of the 8 main technology class-groups in 1990. To control more directly for the local industry composition, in column 5, we include the vector of 1990 local employment shares across the 17 main industries in the Census industrial classification.

To provide further evidence that the instrument is not capturing correlated industry trends across technologically linked cities, column 6 replicates the main 2SLS, with a version of the instrument in which the coefficient of diffusion is set to zero not only when the origin and destination cities coincide, but also when the origin and destination technology classes are the same. In other words, we set $d_{(b,\mu)-(c,\nu)}^c = 0$ whenever either $b = c$ or $\mu = \nu$ (or both). This version of the instrument displays a weaker correlation with observed patenting growth (the F-stat of the first stage drops from 34.7 to 16.9) but the coefficient of the IV regression is robustly positive and slightly larger in magnitude compared to our preferred regression. A related concern is that the instrument captures geographical clustering of innovation that is

\footnote{Hornbeck and Moretti (2018) use a similar variable as an instrument for local productivity shocks.}
Table 3: Patenting growth and economic segregation: robustness

<table>
<thead>
<tr>
<th></th>
<th>ΔIncSegr</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patenting growth</td>
<td></td>
<td>1.30**</td>
<td>1.52**</td>
<td>1.13**</td>
<td>1.62**</td>
<td>1.74***</td>
<td>1.35**</td>
<td>1.22**</td>
<td>2.02***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.61)</td>
<td>(0.76)</td>
<td>(0.57)</td>
<td>(0.67)</td>
<td>(0.65)</td>
<td>(0.62)</td>
<td>(0.62)</td>
<td>(0.67)</td>
</tr>
<tr>
<td>Patenting growth (pre-trend)</td>
<td></td>
<td>-0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.47)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartik-like variable</td>
<td></td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class shares controls (1990)</td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Empl. shares controls (1990)</td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Constr. Instr. (No same class)</td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Constr. Instr. (No neighbor CZs)</td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>State fixed effects</td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Baseline controls</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td># Obs.</td>
<td></td>
<td>722</td>
<td>722</td>
<td>706</td>
<td>706</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>722</td>
</tr>
</tbody>
</table>

First stage estimates

| Predicted patenting growth | 0.59*** | 0.52*** | 0.74*** | 0.65*** | 0.54*** | 0.41*** | 0.40*** | 0.53*** |
|                           | (0.10)  | (0.09)  | (0.16)  | (0.13)  | (0.10)  | (0.09)  | (0.08)  |
| F-stat                   | 34.7     | 30.2     | 21.4     | 26.56    | 29.82    | 16.9    | 19.23    | 41.7     |

Notes: Regressions are weighted by total number of households in 1990. Controls are at 1990 values, with the exception of Import exposure (defined as in Autor et al., 2013). Missing observations in column 3 and 4 are CZs with no patents in the 1990 observation, for which the Bartik-like variable and 1990 shares are not defined. Standard errors clustered at the state level in parenthesis. ***p < 0.01; **p < 0.05; *p < 0.1.

reflected in the citation network. In column 7, we show that results are similar when using a version of the instrument that excludes citations coming from neighboring cities.

We address the concern of changes in legislation and other geographically correlated unobservable factors by introducing state fixed effects in the 2SLS estimation of Equation (4). In this case, we are evaluating changes in segregation resulting from an expansion in innovation activities only through within-state variation. The estimated coefficient, reported in column 8, is positive, significant, and larger in magnitude.

3.5 Inequality and sorting channels

The fact that occupational and educational segregation both display a positive response to innovation shocks suggests that the effect we estimate on income segregation is not simply a by-product of an increase in within-city inequality (inequality channel) but reflects, at least to some extent, a more pronounced sorting of residents along the income dimension (sorting channel).
While the exact magnitude of the inequality and sorting channels on income segregation cannot be precisely pinned down in this empirical framework, it is possible to estimate an upper-bound for the contribution of the inequality channel to the overall effect. Specifically, the extent to which the inequality channel can be responsible for changes in income segregation depends on the initial degree of income segregation in the city. In general, in the absence of any change in sorting patterns – that is, if people are not allowed to relocate – an increase in the dispersion of income across households (measured as an increase in the city-level Gini index) due, for example, to a jump in the college premium, is reflected less than one-to-one into an increase in measured segregation.

To gain intuition on this claim, consider the effect of an increase in income inequality on income segregation in the extreme case of a city that is \textit{perfectly} segregated (i.e., in which residents are perfectly sorted across neighborhoods by income). In this case, in the absence of any change in sorting, an increase in income inequality translates one-to-one into an increase in income segregation. However, in general, cities in the sample are \textit{imperfectly} segregated. In this general case, in the absence of any change in sorting, positive changes in inequality transmit less than one-to-one into changes in segregation. In the opposite extreme case of a city in which the income distribution in each neighborhood is the same (which, incidentally, implies that income segregation is zero), in the absence of any change in sorting, an increase in income inequality would have no impact on income segregation.

In Table 4, we provide a comparison of the impact of patenting growth on income segregation and within-city inequality. Specifically, we estimate Equation (4) using, alternatively, $\Delta IncSegr$ and $\Delta IncIneq$ (the change in the city-level Gini index) as dependent variables. The impact of innovation on inequality (column 2) is smaller than the effect on income segregation (column 1) by more than one third, suggesting that the sorting channel explains \textit{at least} one third, and possibly more, of the response of income segregation to local innovation shocks. This is because, as highlighted in the discussion above, increases in income inequality transmit less than one-to-one to income segregation. According to our estimates, a 10 percentage point increase in patenting growth leads to an increase of 0.091 Gini points in income inequality and 0.130 Gini points in income segregation. This suggests that the sorting channel explains at least 0.039 Gini points of the effect on segregation. Finally, column 3 shows a positive and significant impact (+19.0 percentage points) of patenting growth on the occupation premium, defined as the wage premium earned on the local labor market by workers in knowledge-intensive occupations relative to workers in residual occupations, suggesting that the occupation premium is one of the channels that drives up inequality in response to innovation shocks.

Note that in reality we observe changes in income segregation that are larger than contemporaneous changes in income inequality. This is because, in general, residents respond to innovation (or other) shocks by changing their sorting patterns. Our discussion refers to the transmission of changes in income inequality to changes in income segregation, when people are not allowed to relocate.
### Table 4: Income segregation, inequality, and predicted patenting growth

<table>
<thead>
<tr>
<th></th>
<th>∆IncSegr</th>
<th>∆IncIneq</th>
<th>∆OccPrem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patenting growth</td>
<td>1.30**</td>
<td>0.91**</td>
<td>0.19***</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.40)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Baseline controls</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Estimation</td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
</tr>
<tr>
<td># Obs.</td>
<td>722</td>
<td>722</td>
<td>722</td>
</tr>
</tbody>
</table>

Notes: Regressions are weighted by total number of households in 1990. Controls are at 1990 values, with the exception of Import exposure (defined as in Autor et al., 2013). ∆OccPrem is defined as average earnings by workers in knowledge-intensive occupations divided by the average earnings of workers in residual occupations (averages computed from individual-level data from IPUMS). Standard errors clustered at the state level in parenthesis. ***p < 0.01; **p < 0.05; *p < 0.1.

## 4 Mechanism

The results presented thus far show a robust and economically meaningful causal relationship between an expansion in local innovative activities and an increase in economic segregation in U.S. cities between 1990 and 2010. In this section, we leverage data from the National Establishment Time Series (NETS) to investigate how innovation shocks reshape the internal geography of cities, and provide suggestive evidence for an economic mechanism that can drive this causal relationship.

In line with the extensive body of literature on the centrality of localized idea flows for innovation (e.g., Jaffe et al., 1993, Carlino and Kerr, 2015), we interpret local innovation shocks as an increase in the returns from localized knowledge spillovers, that drives up the incentives of knowledge-intensive firms to cluster in space. We start by showing that, consistently with this interpretation, cities with higher innovation shocks display a more pronounced clustering of knowledge-intensive jobs into neighborhoods characterized by a high initial density of knowledge workers. We then show that this clustering of jobs coincides with a change in the residential choices of knowledge workers, that relocate towards neighborhoods with lower commuting distance from those new employment opportunities, increasing the degree of residential sorting in the city. This initial relocation of knowledge workers can explain part of the effect of innovation shocks on economic segregation. However, we further show that this clustering of residential choices induces responses in both the rental price of housing and measures of local consumption amenities (restaurants, food shops, and fitness centers). If different categories of
residents have heterogeneous valuations of those amenities, or heterogeneous sensitivity of their residential choices to changes in rental prices (as estimated, among others, by Diamond, 2016 and Almagro and Domínguez-Iino, 2021), these responses can significantly amplify the effect of innovation shocks on economic segregation.\footnote{In ongoing work (Berkes and Gaetani, 2021), we develop an extension of Ahlfeldt et al. (2015)'s model that formalizes the feedback link between location of employment and residential choices, and how this link is mediated by rental prices and commuting considerations. The model accounts for the increase in the spatial concentration of knowledge-intensive jobs in response to innovation shocks, and allows to quantify the amplifying effects of endogenous rental prices and supply of residential amenities.}

4.1 Clustering of knowledge-intensive employment

We start by exploring the effect of innovation shocks on the clustering of knowledge-intensive employment within cities. The instrument outlined in Section 3.2.1 identifies shocks to local patenting that result from a higher availability of ideas used as an input for innovation. As proposed and extensively tested in the innovation literature, geographical proximity is a key determinant of knowledge exchange. Hence, when new ideas become available and give rise to innovation opportunities, the incentives of knowledge-intensive firms to co-locate increase and, for the marginal firm, outweigh the costs of co-location. This should result in a more pronounced clustering of knowledge-intensive jobs into neighborhoods that provide better opportunities for knowledge exchange.

To investigate this hypothesis, we use NETS data to construct neighborhood-level measures of local learning externalities and employment in knowledge-intensive and residual occupations (see Appendix B.2 for details on the construction of the employment count by occupation type at the neighborhood level using NETS data). As a neighborhood-level measure of initial learning externalities, we use the density of knowledge-intensive employment in 1990. Specifically, we define learning externalities in neighborhood $n$ as

$$\Lambda_{n,1990} = \frac{W_{n,1990}}{L_n},$$

where $W_{n,1990}$ is the number of workers in knowledge-intensive occupations in 1990, and $L_n$ is the total amount of land (in square km). This formulation mirrors the functional form for productivity externalities in Ahlfeldt et al. (2015) and Tsivanidis (2018), who postulate neighborhood productivity to be a geometric function of the density of employment.\footnote{Ahlfeldt et al. (2015) postulate a more general formulation in which productivity also depends on the density of employment in the surrounding neighborhoods, with the strength of the externality decaying exponentially with distance. However, their unit of analysis (the block) is, on average, significantly smaller than a census tract, and their estimates of the rate of decay imply that externalities decline steeply with distance, so that terms from surrounding neighborhoods are unlikely to contribute significantly to the measure in Equation (8).}

We then test the hypothesis that, in cities that receive higher innovation shocks, knowledge-intensive jobs cluster into neighborhoods with stronger initial learning externalities. Letting
be the change between 1990 and 2010 in the share of workers in \( n \) employed in knowledge-intensive occupations, we estimate the following equation:

\[
\Delta s_n^w = \alpha_{cw} + \beta_{cw} \log(\Lambda_{n,1990}) + \gamma_{cw} \log(\Lambda_{n,1990}) \times \Delta \log(1 + Patents_c) + \epsilon_{nw},
\]

where \( \alpha_{cw} \) is a city fixed effect. A positive sign for the coefficient of the interaction, \( \gamma_{cw} \), suggests that neighborhoods with higher learning externalities, in cities with stronger innovation shocks, have experienced a more pronounced shift towards knowledge-intensive occupations.\(^{25}\)

Columns 1 and 2 of Table 5 display the OLS and 2SLS estimates of Equation (9). We cluster standard errors at the city level, and weight each census tract by the total number of workers in 1990. We multiply coefficients by 100 to interpret them as percentage points. The interaction term has a positive and statistically significant coefficient, that is meaningful in magnitude. Combining the estimates of \( \beta_{cw} \) and \( \gamma_{cw} \) in Column 2, we obtain that in cities at the 95th percentile of the distribution of innovation shocks, a one within-city standard deviation increase in learning externalities is associated with a shift towards knowledge-intensive occupations of 1.01 percentage points, equal to 18.9% of a within-city standard deviation in \( \Delta s_n^w \). The corresponding effect in cities at the 5th percentile of the distribution of innovation shocks is significantly smaller (0.35 percentage points, equal to 6.6% of a within-city standard deviation in \( \Delta s_n^w \)).

4.2 Response in the residential choices of knowledge workers

The estimates of Equation (9) suggest that innovation shocks induce a more pronounced clustering of knowledge-intensive employment into neighborhoods with a high initial density of knowledge-intensive jobs. Because of commuting considerations, changes in the spatial distribution of employment are likely to result into changes in the spatial distribution of residents. In particular, we expect this clustering of jobs to coincide with a more pronounced clustering in the residential choices of knowledge workers. In what follows, we provide suggestive evidence that, consistently with this intuition, stronger innovation shocks induce knowledge workers to relocate their residence to decrease their commuting distance from these employment opportunities. This relocation can explain part of the increase in economic segregation in response to innovation shocks documented in our main results.

\(^{25}\)This is consistent with a model in which neighborhood-level learning externalities \( (\Lambda_{n,1990}) \) and city-level innovation shocks \( (\Delta \log(1 + Patents_c)) \) are complements in the productivity of knowledge-intensive workers, but have no effect on the productivity of workers in residual occupations. In this setting, stronger innovation shocks induce the marginal knowledge worker to relocate to neighborhoods with higher learning externalities. At the same time, the resulting increase in the rental price of office space induces the marginal worker in residual occupations to relocate to neighborhoods with weaker learning externalities. Overall, this leads to a more pronounced increase in the share of knowledge workers \( (s_n^w) \) in neighborhoods with higher learning externalities, in cities with stronger innovation shocks.
Table 5: Clustering of knowledge workers and residents

<table>
<thead>
<tr>
<th></th>
<th>Workers $(\Delta s_n^{w})$</th>
<th>Residents $(\Delta s_n^{r,0−5})$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>log($\Lambda_{n,1990}$)</td>
<td>0.30*** (0.03)</td>
<td>0.28*** (0.03)</td>
</tr>
<tr>
<td></td>
<td>0.26** (0.13)</td>
<td>-0.004 (0.12)</td>
</tr>
<tr>
<td>log($\Lambda_{n,1990}$) $\times$ $\Delta \log(1 + Patents_c)$</td>
<td>0.11*** (0.06)</td>
<td>0.23*** (0.07)</td>
</tr>
<tr>
<td></td>
<td>0.82*** (0.20)</td>
<td>1.90*** (0.26)</td>
</tr>
<tr>
<td>CZ fixed effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Estimation</td>
<td>OLS</td>
<td>2SLS</td>
</tr>
<tr>
<td># Obs.</td>
<td>60,452</td>
<td>60,452</td>
</tr>
<tr>
<td></td>
<td>59,666</td>
<td>59,666</td>
</tr>
<tr>
<td># Clusters</td>
<td>714</td>
<td>714</td>
</tr>
<tr>
<td></td>
<td>714</td>
<td>714</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.05</td>
<td>0.18</td>
</tr>
<tr>
<td>First stage F-stat</td>
<td>136.85</td>
<td>119.7</td>
</tr>
</tbody>
</table>

Notes: $\Delta s_n^{w}$ is defined as the change between 1990 and 2010 in the percentage of workers in census tract $n$ employed in knowledge-intensive occupations. $\Delta s_n^{r,0−5}$ is defined as the change between 1990 and 2010 in the percentage of residents in census tract $n$ employed in knowledge-intensive occupations. Observations are weighted by total number of workers (columns 1 and 2) and residents (columns 3 and 4) in 1990. Standard errors clustered at the CZ level in parenthesis. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

We estimate a specification analogous to Equation (9), but with the dependent variable defined as the change in the percentage of residents in knowledge-intensive occupations residing in neighborhoods within 5 minutes of commuting distance from census tract $n$:

$$\Delta s_{n,0−5}^{r} = \alpha_c^r + \beta^r \log(\Lambda_{n,1990}) + \gamma^r \log(\Lambda_{n,1990}) \times \Delta \log(1 + Patents_c) + \epsilon_n^r, \quad (10)$$

where parameters and variables are defined as in Equation (9), but refer to residents instead of workers. In this case, a positive estimate of the coefficient of the interaction, $\gamma^r$, implies that in cities with stronger innovation shocks residents employed in knowledge-intensive occupations relocate more strongly to neighborhoods at low (0-5 minutes) commuting distance from locations with high initial learning externalities (captured by $\Lambda_{n,1990}$) compared to cities with weaker innovation shocks.

Columns 3 and 4 of Table 5 report the OLS and 2SLS estimates of Equation (10). We cluster standard errors at the city level and weight observations by the total number of residents (in either type of occupation) in 1990.\footnote{Also in this case, we multiply coefficients by 100 to interpret them as percentage points.} The coefficient of the interaction term is estimated to
Figure 7: **Clustering of residents by commuting distance**

Notes: The plot shows 2SLS point estimates and 95% confidence intervals of the coefficient of the interaction of log($\Lambda_{n,1990}$) and $\Delta \log(1 + \text{Patents}_c)$ in specifications analogous to Equation (10). The dependent variable is defined as $\Delta s_{r,n,m}^{1} - m_{2}^{2}$, where $(m_{1} - m_{2})$ is the commuting distance bin indicated on the horizontal axis. Observations are weighted by total number of residents in 1990. Standard errors are clustered at the city level.

be positive and significant, with larger magnitudes – both in absolute terms and relative to the standard deviation of the dependent variable – compared to the effect on the clustering of employment. The estimates in Column 4 imply that, in cities at the 95th percentile of the distribution of innovation shocks, a one within-city standard deviation increase in learning externalities is associated with a 3.33 percentage point increase in the share of residents in knowledge-intensive occupation, equal to 68% of a within-city standard deviation in $\Delta s_{c,t,0-5}^{t}$. The corresponding effect in cities at the 5th percentile of the distribution of innovation shocks is opposite in sign (-2.05 percentage points, equal to a decrease of 41.7% of a within-city standard deviation in $\Delta s_{c,t,0-5}^{t}$).

It is interesting to note that the effect on the clustering of residents is more pronounced than the effect on the clustering of workers. This suggests the existence of a channel that amplifies the effect on residential choices, that we explore in the next subsection.

In Figure 7, we show estimated coefficients of the interaction term when the dependent variable is constructed for different bins of commuting distances. In addition to the baseline (0-5 minutes), we consider bins of census tracts at distance of 5-10 minutes, 10-15 minutes, and so on up to 25-30 minutes from the focal neighborhood $n$. We then estimate the corresponding version
of Equation (10) and report, for each bin of commuting distances, 2SLS estimates of $\gamma^r$ and 95% confidence intervals. The figure shows that the effect of innovation shocks on the relocation of knowledge residents is highly localized and mostly contained in the immediate surroundings of the focal neighborhoods. The effect is the strongest at the lowest commuting distance (0-5 minutes) and declines steeply with space, losing statistical significance for neighborhoods at 15-20 minutes of distance, and turning negative and significant for neighborhoods with distance above 25 minutes.

4.3 Effect on rental price of housing and consumption amenities

The estimates of Equation (10) suggest a significant response in the geographical segmentation of residents by occupation groups, that can explain part of the effect of innovation shocks on measures of economic segregation. As discussed above, the response is more pronounced for the clustering of residents than for the clustering of workers. In this subsection, we provide suggestive evidence for a potential channel of amplification that can account for this difference. We show that the relocation of knowledge residents coincides with a response in the rental price of housing, as well as in measures of local consumption amenities. This response can work as a powerful amplification channel on economic segregation if, as proposed and estimated in the literature (Diamond, 2016; Almagro and Domínguez-Iino, 2021), different categories of individuals (in this case, residents employed in knowledge-intensive and residual occupations) have heterogeneous valuations of those amenities, or heterogeneous sensitivity of their residential choices to changes in the rental price of housing.

We first consider the effect of innovation shocks on the rental price of housing. We estimate a specification analogous to Equation (10), but with the percentage change in the average rental price of housing as dependent variable. The top-left panel of Figure 8 reports the estimated coefficients of the interaction term ($\gamma$) for the average rental price of housing at various bins of commuting distance. The results are strongly in line with those presented in Figure 7. The effect is positive and statistically significant in the surrounding neighborhoods (0-5 minutes of commuting distance), with the point estimate implying that in cities at the 95th percentile of the distribution of innovation shocks, a one within-city standard deviation increase in 1990 knowledge externalities increases the rental price of housing by 10.6%, equal to 51.6% of a within-city standard deviation. The corresponding effect in cities at the 5th percentile of the distribution of innovation shock is opposite in sign (-6.0%, equal to a decrease of 29.2% of a within-city standard deviation). Furthermore, the coefficient of the interaction term steeply declines with space, and becomes negative and significant above 25 minutes of commuting distance.

There are several possible reasons why the rental price of housing might increase (decrease)
in response to an inflow (outflow) of knowledge residents. A natural explanation is the inter-
action of a fixed supply and an increasing demand for residential and office space that drives up
the rental price. A complementary view is that knowledge workers have a higher opportunity
cost of time (Su, 2020), which increases their willingness to pay for housing at lower commuting
distance from neighborhoods with a higher availability of knowledge-intensive jobs. An alter-
native possibility, that we can test using the NETS data, is that the relocation of knowledge
residents prompts a response in the availability of local consumption amenities. If demand for
those amenities comes disproportionately from knowledge workers themselves, either because
of intrinsic differences in preferences (e.g., Baum-Snow and Hartley, 2020; Couture and Hand-
bury, 2020) or because of non-homothetic demand (e.g., Couture et al., 2020), their appearance
can increase the desirability of locations and drive up the rental price of housing.

The top-right and bottom panels of Figure 8 provide suggestive evidence consistent with the
existence of this amplification channel. We plot estimates of the coefficient of the interaction
($\gamma$) in specifications analogous to Equation (10), with the dependent variable defined as

$$\Delta A_{n,m_1-m_2} = 1,000 \times \frac{\Delta K_{n,m_1-m_2}}{R_{n,m_1-m_2,1990}},$$

where $\Delta K_{n,m_1-m_2}$ is the change between 1990 and 2010 in the number of establishments in
category $i$ in census tracts at commuting distance between $m_1$ and $m_2$ minutes from $n$, and
$R_{n,m_1-m_2,1990}$ is the total number of residents in 1990 in the same group of neighborhoods. We
consider three categories of consumption amenities: restaurants, food shops, and fitness centers
(see Appendix B.2 for details on the definitions of these categories).

The estimated coefficients are in line with the view that innovation shocks affect the location
of consumption amenities. In cities with higher innovation shocks, we observe an increase in
the availability of restaurants (top-right panel), food shops (bottom-left panel), and fitness
centers (bottom-right panel), in census tracts at less than 10 minutes of commuting distance
from neighborhoods with higher initial density of knowledge workers. The coefficients of the
interaction in the first bin (0-5 minutes) range from 0.03 per 1,000 residents for food shops to 0.2
per 1,000 residents for restaurants. For comparison, the average within-city standard deviations
in $\Delta A_{n,0-5}$ are equal to 1.36, 0.37, and 0.46 per 1,000 residents for restaurants, food shops,
and fitness centers, respectively. This implies that innovation shocks are a meaningful source
of variation in the evolution of consumption amenities in cities, and the endogenous arrival
of local amenities can work as an amplification channel for the effect of innovation shocks on
economic segregation.
Figure 8: **Rental price of housing and consumption amenities by commuting distance**

<Figure showing the impact of commuting distance on rental prices and amenities.>

**Notes:** The plot shows 2SLS point estimates and 95% confidence intervals of the coefficient of the interaction of log($\Lambda_{n,1990}$) and $\Delta \log(1 + Patents_{i})$ in specifications analogous to Equation (10). The dependent variable is defined as the percentage change in average rent (top-left panel), and $\Delta A_{n,m_1-m_2}$, where $i$ refers to the categories “restaurants” (top-right panel), “food shops” (bottom-left panel), and “fitness centers” (bottom-right panel) in neighborhoods within the commuting distance bin indicated on the horizontal axis. Observations are weighted by total number of residents in 1990. Standard errors are clustered at the city level.

5 Conclusion

In recent decades, policy makers have been increasingly promoting the development of local innovation-based activities, for example by supporting aggressive bids to attract knowledge-intensive firms (Slattery and Zidar, 2020). In many cases, local communities have responded to those attempts with skepticism, fueled by concerns that an expansion of knowledge-intensive jobs would lead to unequally distributed gains and exacerbate urban economic inequalities.

In this paper, we provide the first systematic causal evidence on the role that innovation-based activities play in shaping spatial economic disparities within urban areas. Our IV es-
Estimates suggest that, in the cross-section of U.S. commuting zones, a one standard deviation increase in 1990-2010 patenting growth leads to a 0.65 Gini point increase in income segregation, corresponding to 30.6% of a standard deviation in the 1990-2010 change in income segregation. We provide suggestive evidence of a potential mechanism behind this relationship: An expansion in local innovation possibilities increases the incentives of knowledge-intensive firms to cluster in space to take advantage of localized learning externalities. Because of commuting considerations, this clustering of employment increases residential sorting, with the effect being amplified by the endogenous response of the rental price of housing and consumption amenities.

An important implication of our findings is that, as knowledge-based activities have become the main propulsive force of many local labor markets, the geographical divide between individuals with different levels of income have worsened, possibly exacerbating differences in access to education, health, and consumption amenities. This can contribute to explaining why, in recent decades, workers in the lower portion of the income distribution have witnessed a deterioration across various measures of economic well-being (Autor, 2019; Coile and Duggan, 2019). Understanding the impact of local innovation shocks on the economic outcomes of workers outside of the knowledge economy, and their possible margins of response (e.g., migration, skill upgrade, change in industry, occupation, or labor force status) can be an important avenue for future research.

From a normative perspective, our findings suggest that different local policies can mitigate the undesirable effects of innovation shocks on local spatial inequalities. First, investment in transit infrastructure could make the initial motive for relocation weaker and, at the same time, increase the accessibility of more remote areas in cities. Second, loosening restrictive zoning policies could dampen the sharp response in the rental price of housing, lowering the impact on residential sorting. Third, the amplification generated by the inflow of local consumption amenities could be weakened by redistributive policies that limit the divergence in the quality of amenities across neighborhoods.

The effectiveness of these policies should be evaluated within a framework that can account for general equilibrium interactions. In ongoing work (Berkes and Gaetani, 2021), we develop a quantitative model that extends the framework in Ahlfeldt et al. (2015) to a setting with multiple cities and occupations that we estimate leveraging the same exogenous variation in patenting intensity described in this paper. The quantitative model can be used to disentangle the relative importance of productivity and residential externalities in driving the empirical findings and to perform policy counterfactuals. It can also be used to predict the impact of events such as the arrival of large knowledge-intensive employers on the internal geography of cities, as in Dingel and Tintelnot (2020). This analysis can suggest a path to ensure that the benefits resulting from this secular transformation – that is essential to the economic prosperity of countries and regions – are distributed in a sustainable and inclusive way.
References


——— (2017): *The new urban crisis: How our cities are increasing inequality, deepening segregation, and failing the middle class—and what we can do about it*, Basic Books New York.


## A Additional Tables and Figures

Table A.1: **Summary statistics**

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \log(1 + \text{Patents})$</td>
<td>722</td>
<td>-2.77</td>
<td>3.32</td>
<td>0.18</td>
<td>0.50</td>
</tr>
<tr>
<td>Total patents (1990)</td>
<td>722</td>
<td>0</td>
<td>30,227</td>
<td>6,853</td>
<td>8,610</td>
</tr>
<tr>
<td>Total patents (2010)</td>
<td>722</td>
<td>0</td>
<td>69,719</td>
<td>9,785</td>
<td>13,084</td>
</tr>
<tr>
<td>Average income (1990)</td>
<td>722</td>
<td>18,144</td>
<td>59,698</td>
<td>39,149</td>
<td>8,438</td>
</tr>
<tr>
<td>Number of households (1990)</td>
<td>722</td>
<td>520</td>
<td>4,914,645</td>
<td>1,041,053</td>
<td>1,319,922</td>
</tr>
<tr>
<td>Number of census tracts</td>
<td>722</td>
<td>1</td>
<td>3031</td>
<td>655.4</td>
<td>821.1</td>
</tr>
<tr>
<td>$IncSegr$ (1990)</td>
<td>722</td>
<td>0</td>
<td>27.3</td>
<td>19.2</td>
<td>5.2</td>
</tr>
<tr>
<td>$IncSegr$ (2010)</td>
<td>722</td>
<td>0</td>
<td>30.9</td>
<td>22.0</td>
<td>6.3</td>
</tr>
<tr>
<td>$\Delta IncSegr$</td>
<td>722</td>
<td>-8.3</td>
<td>9.6</td>
<td>2.8</td>
<td>2.1</td>
</tr>
<tr>
<td>$OccSegr$ (1990)</td>
<td>722</td>
<td>0</td>
<td>29.7</td>
<td>21.5</td>
<td>4.2</td>
</tr>
<tr>
<td>$OccSegr$ (2010)</td>
<td>722</td>
<td>0</td>
<td>30.1</td>
<td>21.0</td>
<td>5.4</td>
</tr>
<tr>
<td>$\Delta OccSegr$</td>
<td>722</td>
<td>-25.7</td>
<td>20.2</td>
<td>-0.5</td>
<td>2.7</td>
</tr>
<tr>
<td>$EduSegr$ (1990)</td>
<td>722</td>
<td>0</td>
<td>48.3</td>
<td>32.0</td>
<td>6.4</td>
</tr>
<tr>
<td>$EduSegr$ (2010)</td>
<td>722</td>
<td>0</td>
<td>42.4</td>
<td>30.6</td>
<td>6.7</td>
</tr>
<tr>
<td>$\Delta EduSegr$</td>
<td>722</td>
<td>-19.2</td>
<td>15.6</td>
<td>-1.4</td>
<td>2.8</td>
</tr>
<tr>
<td>Share of college graduates (1990)</td>
<td>722</td>
<td>0.05</td>
<td>0.39</td>
<td>0.20</td>
<td>0.06</td>
</tr>
<tr>
<td>Trade exposure</td>
<td>722</td>
<td>-0.08</td>
<td>25.41</td>
<td>1.14</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Notes:* Summary statistics are weighted by the number of households in the corresponding year.
Figure A.1: Economic segregation: distribution across cities

Notes: Histograms weighted by the number of households in 1990.
### Table A.2: Economic segregation in a selected sample of cities

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Los Angeles, CA</td>
<td>24.0</td>
<td>27.2</td>
<td>+3.3</td>
<td>27.1</td>
<td>29.8</td>
<td>+2.6</td>
<td>35.7</td>
<td>38.3</td>
<td>+2.6</td>
</tr>
<tr>
<td>New York City, NY</td>
<td>27.3</td>
<td>30.9</td>
<td>+3.6</td>
<td>27.0</td>
<td>29.1</td>
<td>+2.1</td>
<td>37.5</td>
<td>36.0</td>
<td>-1.4</td>
</tr>
<tr>
<td>Chicago, IL</td>
<td>25.0</td>
<td>28.8</td>
<td>+3.8</td>
<td>25.2</td>
<td>26.3</td>
<td>+1.1</td>
<td>38.6</td>
<td>37.3</td>
<td>-1.3</td>
</tr>
<tr>
<td>Philadelphia, PA</td>
<td>22.7</td>
<td>29.0</td>
<td>+6.2</td>
<td>24.1</td>
<td>25.3</td>
<td>+1.1</td>
<td>37.4</td>
<td>36.4</td>
<td>-1.0</td>
</tr>
<tr>
<td>Newark, NJ</td>
<td>23.4</td>
<td>28.6</td>
<td>+5.2</td>
<td>23.2</td>
<td>25.4</td>
<td>+2.2</td>
<td>32.5</td>
<td>34.9</td>
<td>+2.4</td>
</tr>
<tr>
<td>Detroit, MI</td>
<td>24.3</td>
<td>27.5</td>
<td>+3.1</td>
<td>25.9</td>
<td>24.9</td>
<td>-1.0</td>
<td>39.5</td>
<td>36.8</td>
<td>-2.6</td>
</tr>
<tr>
<td>Boston, MA</td>
<td>19.4</td>
<td>25.8</td>
<td>+6.4</td>
<td>20.5</td>
<td>23.2</td>
<td>+2.7</td>
<td>31.7</td>
<td>33.6</td>
<td>+1.9</td>
</tr>
<tr>
<td>San Francisco, CA</td>
<td>23.0</td>
<td>28.2</td>
<td>+5.2</td>
<td>23.5</td>
<td>26.5</td>
<td>+2.9</td>
<td>33.1</td>
<td>35.9</td>
<td>+2.7</td>
</tr>
<tr>
<td>Baltimore, MD</td>
<td>22.5</td>
<td>25.4</td>
<td>+2.9</td>
<td>24.0</td>
<td>23.6</td>
<td>-0.4</td>
<td>35.5</td>
<td>34.9</td>
<td>-0.5</td>
</tr>
<tr>
<td>Houston, TX</td>
<td>26.6</td>
<td>28.2</td>
<td>+1.6</td>
<td>28.1</td>
<td>29.0</td>
<td>+0.9</td>
<td>40.9</td>
<td>40.2</td>
<td>-0.7</td>
</tr>
<tr>
<td>Miami, FL</td>
<td>21.0</td>
<td>28.3</td>
<td>+7.3</td>
<td>22.6</td>
<td>22.4</td>
<td>-0.2</td>
<td>29.5</td>
<td>28.8</td>
<td>-0.7</td>
</tr>
<tr>
<td>Bridgeport, CT</td>
<td>22.6</td>
<td>27.0</td>
<td>+4.4</td>
<td>19.7</td>
<td>21.9</td>
<td>+2.2</td>
<td>30.7</td>
<td>32.3</td>
<td>+1.6</td>
</tr>
<tr>
<td>Seattle, WA</td>
<td>17.6</td>
<td>23.9</td>
<td>+6.3</td>
<td>19.2</td>
<td>25.2</td>
<td>+6.0</td>
<td>30.2</td>
<td>34.1</td>
<td>+3.9</td>
</tr>
<tr>
<td>Pittsburgh, PA</td>
<td>21.3</td>
<td>22.2</td>
<td>+0.8</td>
<td>22.4</td>
<td>21.6</td>
<td>-0.8</td>
<td>34.8</td>
<td>32.0</td>
<td>-2.8</td>
</tr>
<tr>
<td>Atlanta, GA</td>
<td>21.9</td>
<td>25.4</td>
<td>+3.5</td>
<td>22.4</td>
<td>22.3</td>
<td>-0.2</td>
<td>35.1</td>
<td>32.2</td>
<td>-2.9</td>
</tr>
</tbody>
</table>

*Notes: The table reports measures of economic segregation for the 15 largest commuting zones in 1990. For clarity, commuting zone names refer to the largest city only.*
Table A.3: Structure of the instrument

<table>
<thead>
<tr>
<th></th>
<th>1995</th>
<th>1996</th>
<th>...</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>...</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>$d_{10}$</td>
<td>$d_{9}$</td>
<td>...</td>
<td>$d_{2}$</td>
<td>$d_{1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>$d_{10}$</td>
<td>...</td>
<td>$d_{3}$</td>
<td>$d_{2}$</td>
<td>$d_{1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$d_{10}$</td>
<td>$d_{9}$</td>
<td>$d_{8}$</td>
<td>$d_{7}$</td>
<td>...</td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$d_{10}$</td>
<td>$d_{9}$</td>
<td>$d_{8}$</td>
<td>...</td>
<td>$d_{1}$</td>
</tr>
</tbody>
</table>

Notes: Structure of the timing of the instrument. To obtain a prediction for patenting activity in 2005 (denoted by $\hat{2005}$), the coefficients of diffusion $d^\tau$ are applied to the actual patenting between 1995 and 2004. To obtain a prediction for patenting activity in 2006, the coefficients of diffusion $d^\tau$ are applied to the actual patenting between 1996 and 2004 for $\tau = 2, \ldots, 10$, and to predicted patenting in 2005 for $\tau = 1$. This process continues for all years up to 2014, where a prediction is obtained by applying the coefficient of diffusion $d^\tau$ to the actual patenting in 2004 for $\tau = 10$, and to predicted patenting between 2005 and 2013 for $\tau = 1, \ldots, 9$. 
Figure A.2: Actual and predicted patenting growth, 1990-2010

Notes: Quantiles of actual patenting growth, 1990-2010

Notes: Quantiles of predicted patenting growth, 1990-2010
### Table A.4: Patenting growth and economic segregation: pre-trend analysis

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A:</strong> ΔIncSegr, 1980-1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted patenting growth</td>
<td>0.16</td>
<td>-0.71</td>
<td>-0.72*</td>
<td>-0.56</td>
<td>-0.57</td>
<td>-0.32</td>
</tr>
<tr>
<td>(0.35)</td>
<td>(0.44)</td>
<td>(0.41)</td>
<td>(0.41)</td>
<td>(0.40)</td>
<td>(0.41)</td>
<td></td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>9.95**</td>
<td>0.92</td>
<td>0.26</td>
<td>1.51</td>
<td>-0.94</td>
<td></td>
</tr>
<tr>
<td>(4.57)</td>
<td>(4.78)</td>
<td>(4.55)</td>
<td>(6.40)</td>
<td>(6.30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log CTs</td>
<td>0.67**</td>
<td>1.58</td>
<td>1.58</td>
<td>1.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.26)</td>
<td>(1.03)</td>
<td>(1.04)</td>
<td>(1.10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log households</td>
<td>-0.90</td>
<td>-0.85</td>
<td>-1.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.07)</td>
<td>(1.09)</td>
<td>(1.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log average income</td>
<td>-0.66</td>
<td>-0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.11)</td>
<td>(2.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Import exposure</td>
<td>-0.46***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td>0.00</td>
<td>0.05</td>
<td>0.17</td>
<td>0.18</td>
<td>0.18</td>
<td>0.22</td>
</tr>
<tr>
<td><strong>Panel B:</strong> ΔOccSegr, 1980-1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted patenting growth</td>
<td>0.69*</td>
<td>0.82*</td>
<td>0.82</td>
<td>0.52</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td>(0.39)</td>
<td>(0.48)</td>
<td>(0.50)</td>
<td>(0.49)</td>
<td>(0.50)</td>
<td>(0.49)</td>
<td></td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>-1.48</td>
<td>0.85</td>
<td>2.11</td>
<td>1.13</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>(3.05)</td>
<td>(4.35)</td>
<td>(4.11)</td>
<td>(6.37)</td>
<td>(5.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log CTs</td>
<td>-0.17</td>
<td>-1.90*</td>
<td>-1.90*</td>
<td>-1.90*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.23)</td>
<td>(1.07)</td>
<td>(1.06)</td>
<td>(1.10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log households</td>
<td>1.71</td>
<td>1.67</td>
<td>1.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.13)</td>
<td>(1.11)</td>
<td>(1.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log average income</td>
<td>0.52</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.02)</td>
<td>(1.91)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Import exposure</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Panel C:</strong> ΔEduSegr, 1980-1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted patenting growth</td>
<td>0.29</td>
<td>0.42</td>
<td>0.43</td>
<td>0.09</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>(0.48)</td>
<td>(0.59)</td>
<td>(0.61)</td>
<td>(0.63)</td>
<td>(0.65)</td>
<td>(0.63)</td>
<td></td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>-1.57</td>
<td>1.32</td>
<td>2.74</td>
<td>5.08</td>
<td>4.93</td>
<td></td>
</tr>
<tr>
<td>(3.58)</td>
<td>(5.27)</td>
<td>(5.32)</td>
<td>(7.46)</td>
<td>(7.30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log CTs</td>
<td>-0.22</td>
<td>-2.16</td>
<td>-2.15</td>
<td>-2.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.28)</td>
<td>(1.36)</td>
<td>(1.37)</td>
<td>(1.38)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log households</td>
<td>1.93</td>
<td>2.02</td>
<td>2.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.34)</td>
<td>(1.33)</td>
<td>(1.35)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log average income</td>
<td>-1.24</td>
<td></td>
<td>-1.21</td>
<td>(1.76)</td>
<td>(1.73)</td>
<td></td>
</tr>
<tr>
<td>Import exposure</td>
<td>-0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td><strong># Obs.</strong></td>
<td>310</td>
<td>310</td>
<td>310</td>
<td>310</td>
<td>310</td>
<td>310</td>
</tr>
</tbody>
</table>

*Notes:* Regressions are weighted by total number of households in 1990. Controls are at 1990 values, with the exception of Import exposure (defined as in Autor et al., 2013). Standard errors clustered at the state level in parenthesis. ***p < 0.01; **p < 0.05; *p < 0.1.
Table A.5: **Patenting growth and economic segregation: unweighted**

<table>
<thead>
<tr>
<th></th>
<th>$\Delta IncSegr$</th>
<th>$\Delta OccSegr$</th>
<th>$\Delta EduSegr$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Patenting growth</td>
<td>0.09 (0.24)</td>
<td>1.04* (0.61)</td>
<td>0.83** (0.37)</td>
</tr>
<tr>
<td></td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td></td>
<td>1.83*** (0.66)</td>
<td>0.80** (0.39)</td>
<td>1.29* (0.73)</td>
</tr>
<tr>
<td>Baseline controls</td>
<td>✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓</td>
</tr>
<tr>
<td>Estimation</td>
<td>OLS</td>
<td>2SLS</td>
<td>OLS</td>
</tr>
<tr>
<td></td>
<td>2SLS</td>
<td>OLS</td>
<td>2SLS</td>
</tr>
<tr>
<td># Obs.</td>
<td>259</td>
<td>259</td>
<td>259</td>
</tr>
<tr>
<td></td>
<td>259</td>
<td>259</td>
<td>259</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.32</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>First stage F-stat</td>
<td>25.95</td>
<td>25.95</td>
<td>25.95</td>
</tr>
</tbody>
</table>

**Notes:** Observations are restricted to CZs with 1990 number of households above 60,000. Controls are at 1990 values, with the exception of Import exposure (defined as in Autor et al., 2013). Standard errors clustered at the state level in parenthesis. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.**
### Table A.6: Patenting growth and economic segregation: Inverse hyperbolic sine transformation

<table>
<thead>
<tr>
<th></th>
<th>$\Delta IncSegr$</th>
<th>$\Delta OccSegr$</th>
<th>$\Delta EduSegr$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Patenting growth (arcsinh)</td>
<td>0.14</td>
<td>1.27**</td>
<td>1.15**</td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.60)</td>
<td>(0.43)</td>
</tr>
<tr>
<td>Baseline controls</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Estimation</td>
<td>OLS</td>
<td>2SLS</td>
<td>OLS</td>
</tr>
<tr>
<td># Obs.</td>
<td>722</td>
<td>722</td>
<td>722</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.41</td>
<td>0.38</td>
<td>0.28</td>
</tr>
<tr>
<td>First stage F-stat</td>
<td>34.73</td>
<td>34.73</td>
<td>34.73</td>
</tr>
</tbody>
</table>

**Notes:** Regressions are weighted by total number of households in 1990. Controls are at 1990 values, with the exception of Import exposure (defined as in Autor et al., 2013). Standard errors clustered at the state level in parenthesis. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.**
B Data description

B.1 Income distribution at the CT level

The NHGIS provides information on yearly household income at the CT level by dividing household into 15 income bins. The lower bounds of each income bin are: 0$, 15,000$, 20,000$, 25,000$, 30,000$, 35,000$, 40,000$, 45,000$, 50,000$, 60,000$, 75,000$, 100,000$, 125,000$, and 150,000$. In order to measure inequality and segregation, we need to approximate the income distribution. For each bracket except for the top one, we assume that all households in that bracket have income equal to the midpoint of the bracket. The top bin is unbounded, with an average that potentially varies substantially across CTs. Our measures will critically depend on the assumptions made to approximate the income distribution within this bracket. The literature has dealt with this issue by either fitting the parameters of an income distribution (usually assumed to be Pareto) or assuming that the average is a fixed percentage above the amount reported in top coded data (usually 40-50% more). These two methods have been subject to several critics.

For our analysis, we design an alternative approach to assign a value to the top bin, and validate our procedure by comparing the resulting segregation index with the corresponding index we obtain by using information on average personal income, that does not require to make any assumption. First, the 5-year 2008-2012 ACS provides CT-level Gini indices using households as basic unit of analysis. For each census tract in 2010, we set the average of the top bin so that the resulting Gini matches the one reported in the ACS. Second, we use the time series of individual-level Gini data at a state level computed by Frank (2009). From there we collect estimates for the Gini index for all the states in 1990 and 2010 and calculate the percentage change. Assuming that the state trends for individual-level Gini are mirrored by the corresponding CT trends for household-level Gini, we set the average income in the top bin so that the percentage change in the Gini index is equal to the one in Frank (2009).

---

27 See for example Autor et al. (2008) and Lemieux (2006).

28 Critics of the former approach have argued that if the underlying distribution is far from the assumed one, a researcher would obtain better results by taking the bin averages. Critics of the latter have pointed to the fact that the assumption of the average income for the last bin is arbitrary. Different methods to deal with binned income data have been reviewed by von Hippel et al. (2016).

29 Note that in 3609 out of 98032 CTs (3.7%) there is no value that allows us to exactly match the Gini reported in the ACS. This might be due to measurement errors or the approximation that all the households earn the average of the income bracket. In this case, our algorithm diverges, either assigning values that are too low (i.e., smaller than 150,000$ which is the lower bound of the top bin) or too high (i.e., bigger than 1,000,000$). When this happens we assign to the CTs in question a default value of 200,000$ which is in line with the 1.4-rule. We experimented with different default values and the main results are robust. Another 908 CTs (or 0.9%) appear in the income data but not in the Gini data. In that case, we try to match the 2010 national Gini.

30 We are not able to match 20,966 (or 21%) of the 1990 CTs with the 2010 data. In this case, we assume
Figure B.3: Income segregation: household VS per capita income

Notes: Bin-scatter plots of the unconditional correlations between income segregation computed using household income and per capita income in 1990, 2010, and change between 1990 and 2010, weighted by total number of households in 1990.

To further validate our procedure, in Figure B.3 we show bin-scatter plots of income segregation in 1990 (left panel), 2010 (center panel), and 1990-2010 change (right panel), using the household income distribution approximated using the procedure described above, and the same measure computed using per capita income at the CT level, which does not require to make arbitrary assumptions on the distribution of income within brackets. The correlation between the two variables is equal to 93% in 1990 and 91% in 2010. The correlation between the 1990-2010 change in the two variables is also remarkably high (44%).

B.2 Other data sources

Residents by occupation

The distribution of residents by occupation at the CT level is constructed as follows. First, from the NHGIS we obtain information on the CT-level distribution of residents according to a coarse definition of occupations, comprising 13 occupations in 1990 and 25 occupations in 2010. Then, using IPUMS, we construct a city-specific crosswalk that maps the coarse definition of occupation into the fine one (386 occupations in 1990 and 454 in 2010). To this end, we exploit the city-specific frequency of each fine occupation code in each coarse category. Occupations are then categorized in two classes: knowledge intensive and non-knowledge intensive. Knowledge intensive occupations are defined according to Florida (2017) definition of “creative class”: “The creative class is made up of workers in occupations spanning computer science and mathematics; architecture and engineering, the life, physical, and social sciences; the arts, design, music, that their Gini is the same as the national one in 1990 (0.43). As we did in 2010, when the algorithm diverges or estimates an implausible value, we assign to the top bin a default value of 200,000$.
Workers by occupation

We assign workers to workplaces using the National Establishment Time Series (NETS). This data set contains information about employment for the near universe of establishments between 1990 and 2010, as well as their location and NAICS code. The latitude and longitude is provided at 5 geographical levels (namely block face, block group, census tract centroid, ZIP code centroid or street level). We allocate workers to each census tract according to the following procedure. First, we assign to a census tract those establishments whose geographical coordinates are provided at a block face, block group or census tract centroid level. Second, we assign the workers of each establishment geo-located at ZIP code level based on the area of the census tracts it contains.\footnote{For example, if a certain ZIP code contains two census tracts that cover 40\% and 60\% of its area, respectively, we assign 40\% of the employment of an establishment assigned to that ZIP code to the first census tract and 60\% to the second one.} We discard all those establishments whose coordinates are missing, more aggregated than the ZIP level, or reported as ZIP codes that do not appear in the NHGIS files (e.g., P.O. boxes). The final data set includes about 10.6 million establishments in 1990 and about 30.6 million establishments in 2010. This procedure gives us an estimate of workers per NAICS at a census tract level.

We use the NETS data in conjunction with the Occupational Employment Statistics (OES) provided by the Bureau of Labor Statistics (BLS) to get an estimate of the occupational distribution of workers in each census tract. The OES reports the percentage of workers active in a certain occupation for each NAICS (SIC90 for 1990) code.\footnote{Since in the 1990s only certain industry codes were reported in each year, we build the crosswalk for 1990 using OES data from 1990 to 1993. Also, since the data are provided for SIC (instead of NAICS) codes, we first build a crosswalk from NAICS to SIC and we then use the appropriate distributions reported in the OES.} Similarly to what we did for the residents, occupations are then assigned to either the knowledge intensive or the non-knowledge intensive category according to the procedure described above.

Local consumption amenities

The establishment count for the three categories of consumption amenities used in Section 4.3 is also computed from NETS. The “Restaurants” category includes establishments from NAICS 722511 (“Full-Service Restaurants”). The “Food Shops” category includes establishments from all the NAICS in 4452 (“Specialty Food Stores”). The “Fitness Centers” category includes establishments from NAICS 71394 (“Fitness and Recreational Sports Centers”).
Commuting times

Commuting times between each pair of CTs are calculated using driving times between the centroids of each census tract. Because of the high number of possible combinations we were unable to use commercial routing services (e.g., Google Maps) and we relied on the Open Source Routing Machine (OSRM).\(^{33}\) The advantage of using the OSRM is that it is possible to run it locally. This allows us to send queries without limits and in parallel. In particular, it was possible to collect data on commuting times for each pair of neighborhoods within each city (for a total of almost 19.4 million pairs) in just a few hours. The disadvantage is that the OSRM does not contain any data on traffic which might underestimate the actual commuting times/costs faced by workers, particularly during rush hour.\(^{34}\)

\(^{33}\)http://project-osrm.org/

\(^{34}\)Note that, because of the lack of traffic data, commuting times are undirected, that is the time necessary to go from A to B and from B to A is the same. The commuting matrices are therefore symmetric and overall contain more than 38.8 million values.
C Citation Network: Demand Pull or Supply Push?

In Section 3.2.1, we developed an instrument for local patenting activity that exploits a prede-termined network of knowledge flows. Our instrument is valid as long as these knowledge links are determined by factors that are orthogonal to the local future economic activity. A possible concern that would invalidate our identification strategy is that the channels captured through the network of citations reflect demand instead of supply links. This would be problematic for the validity of the model, since demand links are likely to be informative about the state of the local economy. To fix ideas, suppose that an IT firm in San Jose supplies innovation to a car manufacturer in Detroit under commission. In this case, our network would record a strong link from San Jose to Detroit, but the associated knowledge flows would violate the orthogonality assumption, since demand from Detroit is likely to be correlated with unobservable factors in Detroit.

The structure of our network and the long time series of patents data can be used to test for the presence of demand-driven links. Formally, we proceed in three steps. First, we use the knowledge network defined in Section 3.2.1 and the observed patenting activity in the period 1985-1994 to get a forward estimate of the patenting activity between 1995 and 2004. Second, we reverse the network and use the patents filed between 2005-2014 to get an upstream estimate of the patenting activity we expect to observe in the period 1995-2004 if the citations were capturing demand links. The reversed network closely mirrors the one defined in (5), but instead of exploiting forward citations, it uses backward citations:

\[ o_{(b,\mu)}^{\tau}(c,\nu) = \begin{cases} \sum_{p \in P(b,\mu)} \frac{\tilde{\text{ShareCit}}_{p}^{\tau}(c,\nu)}{\tilde{\text{TotPat}}_{c,\nu}^{\tau}} & \text{for } \tau \in \{1, \ldots, 10\}, \\ 0 & \text{if } b = c \end{cases} \]

where \( \tilde{\text{ShareCit}}_{p}^{\tau}(c,\nu) \) denotes the share of citations received by patent \( p \) from patents from class \( \nu \) and commuting zone \( c \) filed \( \tau \) years after \( p \), and \( \tilde{\text{TotPat}}_{c,\nu}^{\tau} \) denotes all the potential destination patents in \( (c,\nu) \) at diffusion lag \( \tau \). The coefficient \( o_{(b,\mu)}^{\tau}(c,\nu) \) represents the number of patents of class \( \mu \) in commuting zone \( b \) that we expect to observe upstream if \( \tau \) years later we observe one patent of class \( \nu \) in commuting zone \( c \) downstream. We then compare the two models (the one based on forward citations and the one based on backward citations) to see which one offers the most accurate description of the innovation process. To do this, we follow Acemoglu et al. (2016) and regress the actual 1995-2004 patenting activity on the patenting activity predicted by the two procedures, and controlling for actual patenting in 1985-1994:
Table C.7: Predicting patents with “supply” and “demand” links

<table>
<thead>
<tr>
<th></th>
<th>(\log(1 + P_{95-04,c}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\log(1 + \hat{P}_{95-04,c}^{up}))</td>
<td>0.771*** (0.105)</td>
</tr>
<tr>
<td>(\log(1 + \hat{P}_{95-04,c}^{down}))</td>
<td>-0.153*** (0.045)</td>
</tr>
<tr>
<td>(\log(1 + \hat{P}_{85-94,c}^{actual}))</td>
<td>0.370*** (0.101)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>0.984</th>
</tr>
</thead>
<tbody>
<tr>
<td># Obs.</td>
<td>722</td>
</tr>
</tbody>
</table>

Notes: \(\hat{P}_{95-04,c}^{up}\) is total 1995-2004 patents predicted with the “supply” links of Equation (5). \(\hat{P}_{95-04,c}^{down}\) is total 1995-2004 patents predicted with the “demand” links of Equation (11). Standard errors clustered at the state level in parenthesis. *** \(p < 0.01\); ** \(p < 0.05\); * \(p < 0.1\).

\[
\log(1 + P_{95-04,c}^{actual}) = \alpha + \beta \log(1 + \hat{P}_{95-04,c}^{up}) + \gamma \log(1 + \hat{P}_{95-04,c}^{down}) + \delta \log(1 + P_{85-94,c}^{actual}) + \epsilon_c,
\]

where \(\hat{P}_{95-04,c}^{up}\) is patenting activity predicted by the model in Equation (5), whereas \(\hat{P}_{95-04,c}^{down}\) is the patenting activity predicted by the model in Equation (11). The results, reported in Table C.7, show that only “supply” links (\(\hat{P}_{95-04,c}^{up}\)) have a strong predictive power, while “demand” links (\(\hat{P}_{95-04,c}^{down}\)) have a small negative impact on actual patenting. The sign and magnitude of the estimates are consistent with the ones obtained by Acemoglu et al. (2016), who consider an analogous setting but use the network of citations to predict patenting growth across technological fields and do not consider its geographical dimension.
Stability of citation network

In this Section, we perform a comparison between the citation network in the early sample and its counterpart in the late sample to verify that the channels of knowledge diffusion inferred from the citations patterns are, at least to some extent, stable over time. We do this in three steps. First, we build the network of citations and compute the coefficients of diffusion separately for the two samples (1975-1994 and 1995-2014). For each $\tau = 1, ..., 10$, we take the difference of the two adjacency matrices and calculate its Frobenius norm as follows:

$$
\text{real}_\tau = \| \mathbf{D}_{75-94}^{\tau} - \mathbf{D}_{95-14}^{\tau} \|_2.
$$

Second, for each year between 1975 and 2014, we reshuffle all the patents filed in that year under the constraint that after the reshuffling each commuting zone is assigned the same amount of patents as in the real dataset. We repeat the same exercise performed in the first step for this new sample of patents and calculate

$$
\text{random}_\tau = \| \tilde{\mathbf{D}}_{75-94}^{\tau} - \tilde{\mathbf{D}}_{95-14}^{\tau} \|_2,
$$

where $\tilde{\mathbf{D}}_{75-94}^{\tau}$ and $\tilde{\mathbf{D}}_{95-14}^{\tau}$ are the citation networks built using the reshuffled patents.

Finally, we calculate the percentage difference between $\text{real}_\tau$ and $\text{random}_\tau$ for each $\tau = 1, ..., 10$. This number captures the distance between the two actual networks (75-94 and 95-14) compared to two networks that, while maintaining the same structure and properties of the original ones, are by construction uninformative of each other. A positive value indicates that the networks built using the actual data are more similar than the reshuffled ones. Figure D.1 plots the percentage difference for all the values of $\tau$ together with the 95% confidence interval we obtained by repeating this procedure 50 times. The difference of the random networks is more than 40% larger than the one obtained with the actual networks for the first lag and it declines for larger lags. The decline implies that the more years pass after a new idea is generated the less the citation patterns are distinguishable from links that are distributed across cities at random. This result is intuitive. With time a new technology becomes widespread and is embedded in patents produced in areas that do not have any direct link with the origin city-class pair.
Figure D.1: Proximity of 75-94 and 95-14 random relative to actual citation networks

Notes: Percentage difference of the Frobenius norms \( \text{random}_\tau \) and \( \text{real}_\tau \) for \( \tau = 1, \ldots, 10 \).