

Income Segregation and the Rise of the Knowledge Economy[†]

By ENRICO BERKES AND RUBEN GAETANI*

We analyze the effect of an increase in knowledge-intensive activities on spatial inequality in US cities. We leverage a predetermined network of patent citations to instrument for local innovation trends. Between 1990 and 2010, a one-standard-deviation increase in patent growth increases income segregation by 0.65 Gini points, corresponding to 0.31 standard deviations of the over-time change in income segregation. This effect mainly arises from the sorting of residents by 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. (JEL D31, O31, O33, O34, R23, R32)

Over the past 40 years, economic activities that rely on nonmanual, nonroutine technical skills, scientific knowledge, and intellectual creativity have become the main engines of economic prosperity in advanced countries (Powell and Snellman 2004). Since 1975 the share of added value 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, the number of scientific publications, the ratio of

* Berkes: The Ohio State University (email: berkes.8@osu.edu); Gaetani: University of Toronto (email: ruben.gaetani@utoronto.ca). Neale Mahoney was coeditor for this article. We thank Matthias Doepke, Benjamin Jones, Bruce Weinberg, and two anonymous referees for their invaluable guidance. We also thank for their comments Treb Allen, Nathaniel Baum-Snow, Davide Coluccia, Jonathan Dingel, Richard Florida, Alberto Galasso, Mitchell Hoffman, Lorenz Kueng, Marti Mestieri, Peter Nencka, and Matthew Notowidigdo; our discussants, Ed Glaeser, Henry Overman, and Jeffrey Zabel; and seminar participants at Northwestern University, University of Toronto, McMaster University, Ryerson University, EIEF, Bocconi University, Centre de Recerca en Economia Internacional, INSEAD, Toulouse School of Economics, The Ohio State University, Carnegie Mellon University, Federal Reserve Bank of Chicago, 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 Innovaton (Cambridge), Duke Fuqua Strategy Conference, Conference on Urban and Regional Economics (Federal Reserve Bank of Philadelphia), 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 this paper are solely our responsibility.

[†] Go to <https://doi.org/10.1257/app.20210074> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

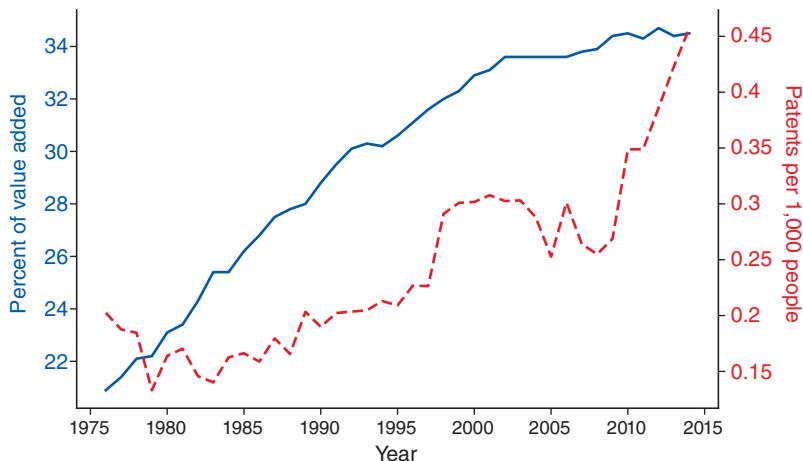


FIGURE 1. MEASURES OF KNOWLEDGE-INTENSIVE ACTIVITIES

Notes: The blue line (left axis) is the contribution to US gross domestic product (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, and arts, entertainment, and recreation (data from Bureau of Economic Analysis 2018). The dashed red line (right axis) is the number of patents per 1,000 people issued to US inventors by the USPTO.

intangibles to assets, and the share of workers employed in research and development (R&D) and creative sectors (Florida 2002). Proposed explanations for this structural shift include globalization, the 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 has examined 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 specializing 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 over the last few decades—namely, the marked increase in 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, which increased from 42.8 to 47.0 Gini points during the same period according to census bureau estimates. To address this question empirically, we exploit cities’ heterogeneous exposure to the rise of the knowledge economy and investigate the relationship between local expansion in knowledge-based activities,

as measured by patenting growth, and changes in economic segregation between 1990 and 2010.¹

Theoretically, there are several reasons to presume the existence of this causal link. First, innovation and other creative jobs depend, crucially, on localized knowledge spillovers, as an extensive literature has previously demonstrated (Glaeser et al. 1992; Jaffe, Trajtenberg, and Henderson 1993; Carlino and Kerr 2015). This implies that an increase in the returns to accessing new ideas promotes geographic clustering. Second, knowledge economy workers tend to be disproportionately sensitive to local amenities, such as school quality, which reinforces the incentives for geographic 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 due to potential reverse causation and the presence of unobservable factors such as local financial or housing shocks that jointly affect the urban environment and a geographic area's ability 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 georeferenced US patents for 1975–2014. We propose that the production of local innovation is at least in part the result of the diffusion and recombination of ideas generated elsewhere in the economy that propagate across geographic areas and technology classes through predetermined channels that are persistent over time. To measure the strength of these channels, we leverage the observed citation network in the early sample (1975–1994) and construct for each city a measure of local exposure to the emergence of external ideas generated in other places. We combine this measure with observed patenting in other locations and technology classes to predict 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 address possible validity concerns. In particular, among other tests, we show that instrumented patenting growth is not correlated with pre-trends in segregation and that our results are robust to controlling for local industrial composition and local exposure to nationwide 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 an increase of one cross-city standard deviation in patenting growth causes a 0.65 Gini point increase in income segregation, corresponding to 0.31 standard deviations of the change in income segregation

¹ A rich literature has shown that economic segregation has a first-order impact on several policy-relevant local outcomes including schooling choices (Katz, Kling, and Liebman 2001; Baum-Snow and Lutz 2011), health (Acevedo-Garcia et al. 2003; Alexander and Currie 2017), and intergenerational mobility (Chetty, Hendren, and Katz 2016; Fogli and Guerrieri 2019).

² Our procedure extends the model proposed by Acemoglu, Akcigit, and Kerr (2016) to a setting with multiple geographic locations. The strategy we propose is general and can be applied to other contexts in which channels of knowledge diffusion are measurable.

during 1990–2010. We find consistent results when considering alternative measures for 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 geographic sorting of households along the income dimension (sorting channel).

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, Walls and Associates 2014) to analyze changes in the locations 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 collocate. As proposed and extensively tested in the innovation literature, geographic proximity is a key determinant of idea transmission, and learning externalities rapidly decline with distance (Jaffe, Trajtenberg, and Henderson 1993; Carlino and Kerr 2015; Catalini 2018). For this reason, when new knowledge becomes available and gives rise to innovation opportunities, firms' incentives to collocate increase and, for the marginal firm, outweigh congestion costs. Consistent with this interpretation, we document that larger innovation shocks are associated with 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 census tracts that experience an inflow of knowledge-intensive jobs, presumably to reduce their commuting distance from these new employment opportunities. This effect declines sharply with space and even reverses for neighborhoods located more than 20 minutes away in terms of commuting distance. Additionally, we show that the rental price of housing and local consumption amenities (as measured by the number of restaurants, food shops, and fitness centers) display an analogous spatial response. To the extent that knowledge workers have a higher valuation of those amenities or a lower sensitivity in their residential choices to changes in rent, this response can contribute to 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 US cities since 1970. More recently, Reardon and Bischo (2016) document that this trend continued to a lesser extent until very recently and is correlated with the increase in income inequality. Baum-Snow and Pavan (2013) show that larger cities experienced larger increases in inequality and interpret this relationship as evidence of a skill-biased change in agglomeration economies. Eckert, Ganapati, and Walsh (2020) document that a subset of industries

(“skilled scalable services”) are largely responsible for the simultaneous increase in the education and urban wage premia since 1980. Our paper is closely related to Diamond (2016) and Rossi-Hansberg, Sarte, and Schwartzman (2019), who study the growing geographic segmentation of individuals by skill across US cities, while our focus is on the determinants of sorting within cities.

A growing literature in urban economics investigates the determinants of recent changes in the internal structures of cities. Couture et al. (2020) develop a model with nonhomothetic preferences in which an increase in income inequality generates demand for high-quality amenities, leading to the gentrification of downtown neighborhoods and increasing welfare disparities between high- and low-income households. Couture and Handbury (2020) study the drivers of downtown revivals since 2000 and find evolving preferences for downtown amenities among young college graduates to be the largest contributor. Su (2022) argues that the rising value of time for skilled workers increases their demand for central locations, contributing to downtown gentrification. Our paper provides evidence that the emergence of knowledge-based activities affects the geographic distribution of residents and jobs in cities, with changes in the location of consumption amenities amplifying the increase in the sorting of residents. This finding contributes to an expanding literature that studies the geography of consumption activities in cities. Recent contributions to this literature include Davis et al. (2019), who document the determinants of consumption segregation in New York City; 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 amenities in a setting that allows for heterogeneous preferences among agents.

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) relate the increase in urban segregation in US metro areas to the expansion of high-technology industry jobs. In our paper, we provide causal evidence that supports their interpretation.

The remainder of the paper is organized as follows: Section I presents the data sources and describes the measures of economic segregation used throughout the paper. Section II describes the main correlations and the IV strategy and discusses the instrument and robustness results. Section III outlines and tests an economic mechanism that might explain the main results. Section IV offers avenues for future research and concludes.

I. Data and Measurement

We combine data on innovation activities, captured by patents, with social and economic indicators from the US 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 these terms interchangeably throughout the manuscript. Commuting zones are defined with respect to actual commuting flows in

the US and, contrary to metropolitan statistical areas, constitute a complete partition of the country.³ Since our objective is to assess how innovation activities affect the concentrations of residents and workers within local labor markets, and since these two aspects are intrinsically connected through commuting decisions, commuting zones are the natural unit of geographic aggregation for this analysis.

A. Patent Data

We use city-level patenting as a proxy for local knowledge intensity. Patent data are collected from the USPTO Bulk Data Storage System (United States Patent and Trademark Office 2015). We parse the text of all 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,⁴ forward and backward citations, and the 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 of which 2,634,606 are located in the United States.⁵

B. 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 each decennial census until 2000 and the ACS from 2005 onward.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 a cross-neighborhood, within-city Gini index and is meant to capture income dispersion in a given city once the variation of income in each neighborhood has been removed. Formally, letting N_c be the number of neighborhoods in city c and indexing neighborhoods by n in ascending order from the lowest average household income to the highest, income segregation is defined as

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

³We construct the commuting zones shapefile (Berkes and Gaetani 2023) using the 1990 definition of commuting zones provided by the Integrated Public Use Microdata Series (IPUMS, Ruggles et al. 2021).

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

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

⁶We use 2008–2012 ACS averages to generate economic and demographic variables for the 2010 observations. CT-level income data divide households into 15 brackets with lower bounds of \$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 CTs. The literature has approached this issue in various ways, each with its own advantages and disadvantages. Online Appendix B.1 discusses these approaches and provides a detailed description of the procedure we use to approximate the income distribution.

where $h_{n,c}$ and H_c denote the number of households and $x_{n,c}$ and X_c denote the total income of neighborhood n and city c , respectively. It is easy to show that in an extreme case, in which the average income of each neighborhood is the same, $IncSegr_c$ is equal to zero. Conversely, if households are perfectly sorted across neighborhoods, income segregation is equal to the standard within-city Gini index.

We measure segregation by occupation and education using an index of dissimilarity,⁷ which captures the asymmetry across neighborhoods in terms of a group of individuals' prevalence. In our case we define these groups as residents who are employed in knowledge-intensive occupations (for occupational segregation) and residents with at least a four-year college degree (for educational segregation). Formally, we define occupational segregation as

$$(2) \quad OccSegr_c = 100 \times \sum_{n=1}^{N_c} \frac{1}{2} \left| \frac{r_{n,c}^1}{R_c^1} - \frac{r_{n,c}^0}{R_c^0} \right|,$$

where $r_{n,c}^1$ and R_c^1 denote the number of residents employed in knowledge-intensive occupations (knowledge workers) in neighborhood n and city c , respectively, and $r_{n,c}^0$ and R_c^0 denote the corresponding number of residents employed in residual occupations. Online 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

$$(3) \quad EduSegr_c = 100 \times \sum_{n=1}^{N_c} \frac{1}{2} \left| \frac{r_{n,c}^H}{R_c^H} - \frac{r_{n,c}^L}{R_c^L} \right|,$$

where the superscript H denotes educational attainment of at least a four-year college degree and L represents any educational attainment below a four-year college degree.

Online 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 13.2 percent 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 percent and 10.1 percent of their nationwide averages in 1990 (21.5 and 32.0, respectively). Online Appendix Table A.2 reports the corresponding statistics for the 15 largest commuting zones. Between 1990 and 2010, some of the major cities (including San Francisco, Boston, and Seattle) experienced an increase in income segregation greater than five Gini points, while other major urban areas (including Pittsburgh and Houston) recorded considerably smaller increases of less than two Gini points.

⁷ See White (1983) for a discussion of the pros and cons of using an index of dissimilarity as a measure of segregation.

C. Other Data Sources

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

Data on employment and local amenities are compiled from the NETS, which provides data on employment, geographic locations, and industry codes for close to the universe of US establishments over the 1990–2015 period.⁸

Bilateral commuting times across census tracts are computed using the Open Source Routing Machine (Luxen and Vetter 2011). This routing engine allows us to compute travel time by car for each pair of neighborhoods in each city for a total of almost 19.4 million pairs.⁹

Finally, the rental price of housing is computed from the NCGIS as the average rent per room.¹⁰ Online Appendix Table A.1 and online Appendix B provide summary statistics and further details on the construction of the main variables.

D. 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 years (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 geographic areas and technology classes and to measure innovation for the 1990 observations. The late sample (1995–2014) is 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 observations. The time structure of the data is illustrated in Figure 2.

II. 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 these phenomena and investigate its features. We then provide evidence for a plausible mechanism underlying this relationship.

⁸ More details on the NETS data and the procedure to assign each establishment to a census tract can be found in online Appendix B.2.

⁹ The data supplement (Berkes and Gaetani 2023) includes the resulting datasets that assign to each neighborhood the set of CTs that lie within various bins of commuting time (e.g., five minutes or less, between five and ten minutes, etc.).

¹⁰ Specifically, average rent per room is computed as the aggregate gross rent for renter-occupied housing units paying cash rent divided by the aggregate number of renter-occupied rooms.

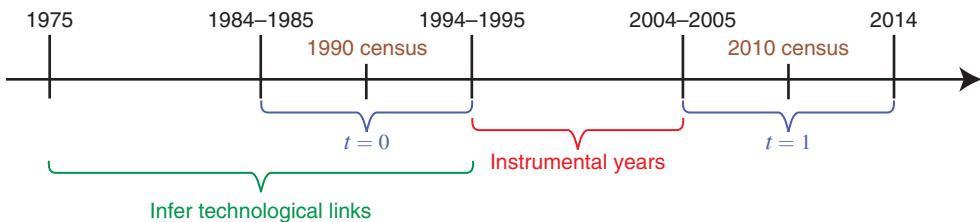


FIGURE 2. DATA TIMELINE

Notes: The $t = 0$ observation corresponds to 1985–1994 data for patenting and the 1990 census for the economic and demographic variables. The $t = 1$ observation corresponds to 2005–2014 patenting data for patenting and the 2008–2012 ACS for the economic and demographic variables.

The empirical model that relates changes in economic segregation and patenting growth at the city level between 1990 and 2010 is specified as follows:

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

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, β , with the outcome variable, ΔY_c , representing the change in each of the three measures of economic segregation defined in Section IB. In the regressions, we weight cities by the number of households in 1990.¹¹ To avoid dropping observations with zero patents in either 1990 or 2010, we adopt the convention of taking the logarithm of one plus the total number of patents.¹²

A. Correlations and Ordinary Least Squares (OLS) Results

The upper-left panel of Figure 3 shows a scatterplot of the unconditional correlation between changes in income segregation and the growth rate in the total number of 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 upper-right and lower-left panels of Figure 3 show analogous positive correlations for the measures of occupational and educational segregation.

¹¹ In online Appendix Table A.6, we report ordinary least squares (OLS) and two-stage least squares (2SLS) results for regressions that are unweighted but only include cities with more than 60,000 households in 1990. The resulting sample includes 259 cities and accounts for 88.8 percent of all US households in 1990. The results are strongly consistent.

¹² In Table A.7 we show that results are robust to using an inverse hyperbolic sine (arcsinh) transformation. Because instances of zero patent counts are concentrated in scarcely populated areas and the observations are weighted by population, we obtain virtually identical results when adopting other strategies used in the literature (e.g., including dummy variables for zeros or calculating growth rates through the midpoint method). Since patenting activity measures are based on ten-year totals, only 16 CZs had no patenting activity in either 1990 or 2010. The total number of households in those cities was about 0.04 percent of all US households in 1990.

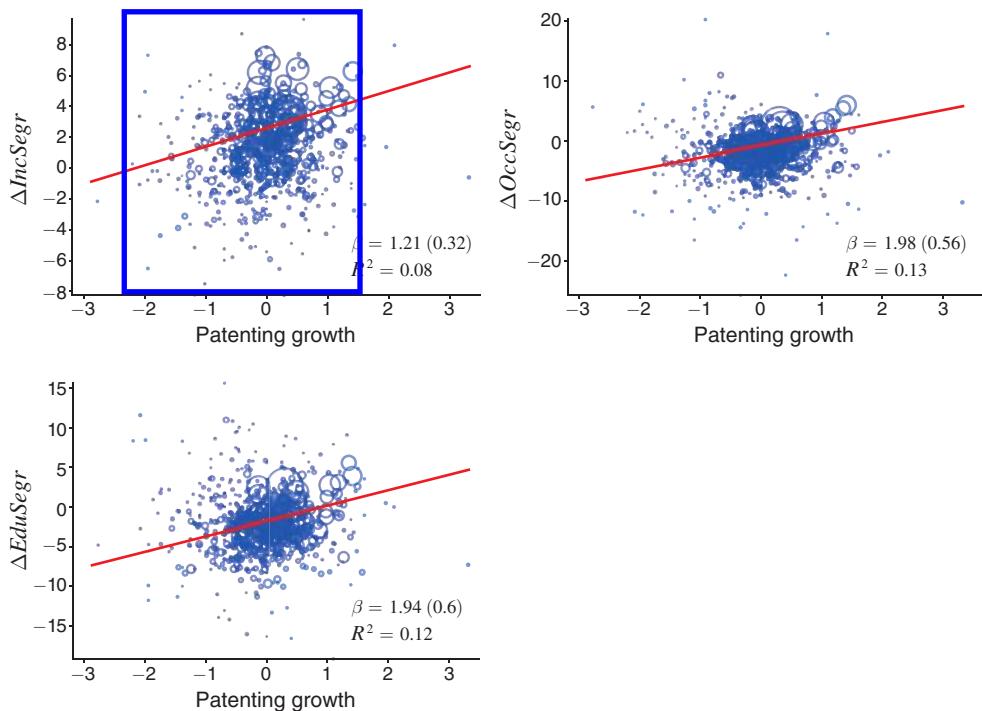


FIGURE 3. PATENTING GROWTH AND ECONOMIC SEGREGATION: UNCONDITIONAL CORRELATIONS

Notes: Scatterplots show the unconditional correlations between growth in patenting and change in income segregation (upper-left panel), occupational segregation (upper-right panel), and educational segregation (lower-left panel) between 1990 and 2010. Circles and regression lines are weighted by the total number of households in 1990.

Table 1 reports the regression results for each of the three measures of segregation after controlling for potential confounding factors including the share of individuals with at least a four-year college degree, the log number of census tracts, the log number of households, and the log average income in 1990. These controls account for the fact that an increase in innovation activities and income segregation is likely to be correlated with the overall level of human capital, the city size, and 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 imports from China developed by Autor, Dorn, and Hanson (2013a).

The inclusion of the full set of controls significantly dampens the estimated effect of patenting growth on all three measures of segregation. While the coefficients remain significant for occupational and educational segregation (panels B and C), they lose statistical significance in the case of income segregation (panel A). This loss of significance is due to the inclusion of the control for the local share

TABLE 1—PATENTING GROWTH AND ECONOMIC SEGREGATION: OLS RESULTS

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. $\Delta IncSegr$</i>						
Patenting growth	1.21 (0.32)	0.27 (0.36)	0.40 (0.31)	0.11 (0.32)	0.16 (0.30)	0.15 (0.30)
Share of college graduates		16.60 (3.87)	7.23 (4.26)	8.27 (3.96)	-0.22 (4.63)	-0.01 (4.55)
log CTs			0.55 (0.16)	-2.35 (1.09)	-2.45 (1.17)	-2.43 (1.20)
log households				2.80 (1.11)	2.54 (1.25)	2.53 (1.27)
log average income					4.55 (1.34)	4.52 (1.34)
Import exposure						0.07 (0.09)
R^2	0.08	0.25	0.33	0.38	0.40	0.41
<i>Panel B. $\Delta OccSegr$</i>						
Patenting growth	1.98 (0.56)	1.00 (0.56)	1.20 (0.51)	1.16 (0.46)	1.21 (0.44)	1.21 (0.44)
Share of college graduates		17.47 (4.15)	2.19 (3.59)	2.34 (3.57)	-5.98 (4.96)	-6.08 (4.99)
log CTs			0.90 (0.12)	0.48 (0.74)	0.38 (0.78)	0.38 (0.77)
log households				0.41 (0.71)	0.15 (0.70)	0.16 (0.70)
log average income					4.47 (2.27)	4.48 (2.28)
Import exposure						-0.03 (0.12)
R^2	0.13	0.24	0.37	0.37	0.38	0.38
<i>Panel C. $\Delta EduSegr$</i>						
Patenting growth	1.94 (0.60)	1.16 (0.56)	1.27 (0.55)	1.15 (0.48)	1.24 (0.43)	1.21 (0.42)
Share of college graduates		13.97 (3.83)	5.36 (3.91)	5.80 (3.77)	-9.96 (6.03)	-9.32 (5.87)
log CTs			0.51 (0.22)	-0.72 (1.20)	-0.90 (1.19)	-0.85 (1.19)
log households				1.18 (1.19)	0.70 (1.11)	0.68 (1.12)
log average income					8.45 (2.60)	8.36 (2.53)
Import exposure						0.21 (0.14)
R^2	0.12	0.18	0.22	0.22	0.28	0.28
Observations	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 (provided by Autor, Dorn, and Hanson 2013b). Standard errors clustered at the state level are in parentheses.

of college graduates, possibly reflecting the fact that 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), among others. Since income segregation is intrinsically connected to inequality, as

we explore in depth in Section III.5, and patenting has grown more in cities with a higher share of college graduates, it is not surprising that the OLS estimate of β decreases once we control for these city-level characteristics.

B. 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 simultaneously affect 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 the 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 a knowledge-based economy in local US labor markets.

In this section we propose an instrument for local growth in innovative activities. The intuition behind the instrument is that local patenting is, at least in part, the result of the recombination of ideas generated elsewhere in the economy. These ideas propagate across geographic 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.¹³ 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 during the 2005–2014 period.

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 three 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–Computers and 3 more (out of 5) from Chicago–Metallurgy. Under the assumption that external ideas are perfectly substitutable inputs for local innovation, these observed citation links suggest that we need 100 new patents in San Francisco–Computers or 5 new patents in Chicago–Metallurgy (the origin city-class pair) to generate one-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 0.5/100 and 0.5/5 for San Francisco–Computers and Chicago–Metallurgy,

¹³The main departure from a traditional input-output production model is that ideas are nonrival inputs in our case. As a result, the sum of all of the inputs that appear in the production of new patents is generally greater than the overall amount of available inputs.

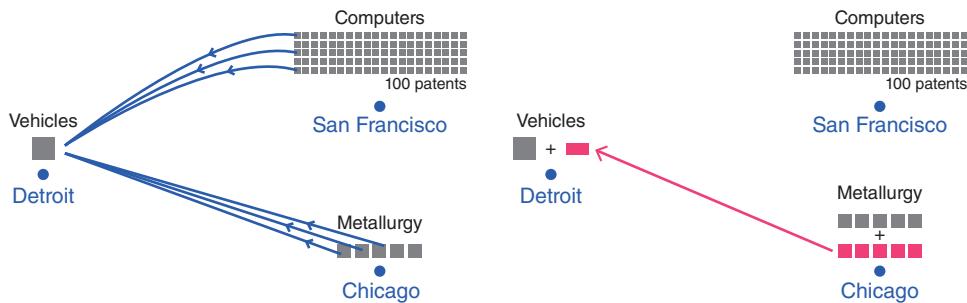


FIGURE 4. PREDICTING PATENTING: AN ILLUSTRATIVE EXAMPLE

Notes. The figure shows an 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 0.5/100 and 0.5/5, respectively. Right panel: In the late sample, Chicago–Metallurgy produces five additional patents. Our procedure predicts an additional half patent in Detroit–Vehicles.

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 five additional patents, we predict an additional one-half of a grant ($0.5/5 \times 5$) in Detroit–Vehicles.

This example illustrates the basic intuition underlying the instrument. In our implementation we generalize this procedure by considering the time lag between the cited and citing patent.¹⁴ In the following subsection we formalize this intuition and describe our procedure in detail.

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 ν (the destination city-class pair), we calculate the share of citations given to patents from city b of class μ (the origin city-class pair) filed τ years before. We interpret the sum of all of these shares, normalized by the total number of patents in (b, μ) , 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, ν) by one patent filed τ years before in (b, μ) .¹⁵

¹⁴Introducing the time lag between cited and citing patents allows us to account for the diffusion of ideas over time through direct as well as indirect links in the network. Moreover, it allows us to account for heterogeneity across technology classes in the speed of diffusion of new ideas.

¹⁵Note 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 that knowledge flows from the origin (cited) to the destination (citing) patent.

To compute the coefficients of diffusion, we consider each patent filed between 1985 and 1994 as a potential destination patent and those filed between one and ten years before as potential origin patents.¹⁶ Mathematically,

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

where $\mathcal{P}(c,\nu)$ is the set of patents in the destination city-class (c,ν) 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 $p \in \mathcal{P}(c,\nu)$ to origin patents in (b,μ) filed τ years before ($ShareCit_{(b,\mu) \rightarrow p}^{\tau}$). The denominator ($TotPat_{b,\mu}^{\tau}$) is the total number of potential origin patents in (b,μ) 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 pair in $2005 - \tau$ by the relevant coefficients for $\tau = 1, \dots, 10$ and take the sum of the resulting numbers:

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

The term κ_{2005} is a rescaling factor that ensures that the total number of patents we estimate nationwide is the same as the number that we observe in the data. The same strategy is used to predict patenting in 2006, with the exception that the coefficients at $\tau = 1$ are applied to predicted patents in 2005 instead of actual patents.¹⁷ We do this to avoid endogeneity concerns that might arise when using contemporaneous patenting. This process is repeated sequentially to predict patenting for every year up to 2014. Online 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).¹⁸ In the remainder of this paper, we refer to this measure as predicted patenting growth or, interchangeably, as an innovation shock.

¹⁶This implies that all patents filed between 1975 and 1993 are potential origin patents. Note that in general, destination patents can also be origin patents.

¹⁷The role of κ_{2005} is now evident: it prevents the national predicted number of patents in later years from being altered by the use of predicted patents alongside actual patents.

¹⁸Recall that patenting in the 1990 (2010) observations is defined as the sum of actual or predicted patents in the years 1985–1994 (2005–2014).

Conditions for Validity.—To identify the causal effect of innovative activities on economic segregation, the usual conditions for instrument validity must hold.¹⁹ First, the instrument must have predictive power on actual patenting growth between 1990 and 2010. We discuss this point extensively in the next subsection; we show that the network of diffusion inferred in the early sample is in fact persistent and successfully predicts 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 the IV settings used in the economic geography literature (e.g., Baum-Snow 2007; Duranton and Turner 2012; and Agrawal, Galasso, and Oettl 2017), the orthogonality requirement in our analysis must hold conditionally on the set of local characteristics that might affect the formation of knowledge linkages in the early sample and that might be correlated with the evolution of local segregation since 1990. In particular, it is plausible that cities with higher 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, have experienced 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 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 itself be an outcome of local shocks in the destination city-class, leading to a violation of the orthogonality condition. To rule out this possibility, in online Appendix C we perform a test in the spirit of Acemoglu, Akcigit, and Kerr (2016) in which we exploit the asymmetric nature of the citation network and predict patenting in 1995–2004 alternatively using “supply” links from 1985–1994 (with backward citations) and “demand” links from 2005–2014 (with forward citations). We show that, conditional on supply links, demand links do not have predictive power for 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.

First-Stage Results.—The first condition for the instrument to be valid is that predicted patenting growth, as obtained through the procedure previously outlined, is

¹⁹Note 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, Sorkin, and Swift (2020) and Borusyak, Hull, and Jaravel (2018).

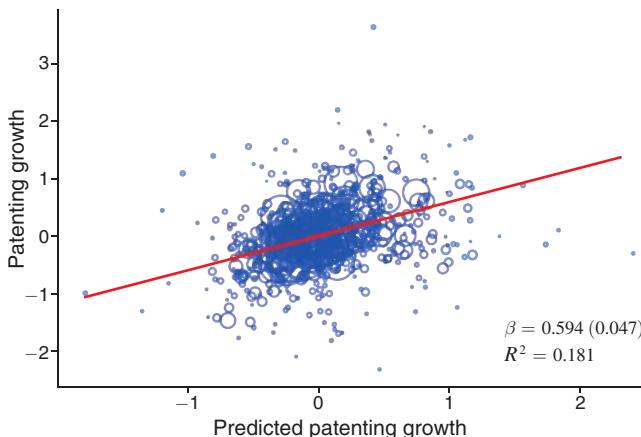


FIGURE 5. FIRST-STAGE SCATTERPLOT

Notes: The figure shows 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 the regression line are weighted by the total number of households in 1990.

correlated with actual patenting growth. This condition requires that the network of knowledge diffusion inferred from the citation patterns is at least to some extent stable over time so that the diffusion links observed in the early period are informative of the links in the late period. This fact can be tested directly by comparing the citation networks in the two samples. In online 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 so that the total number of grants for each city-year remains unchanged. The results indicate that the two observed networks are significantly closer to each other than the simulated ones, suggesting that the network retains its structure over time.

Figure 5 shows a scatterplot of the first-stage relationship between the predicted and actual growth rate of patenting. We plot the residuals of the 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. Online Appendix Figure A.2 visually compares actual and predicted patenting growth at the city level on a map of the United States.

Pre-trend Analysis.—The second condition for instrument validity is that cities that experienced different innovation shocks, as captured by the instrument, should not have followed systematically different trajectories in terms of economic segregation in the absence of shocks. We provide evidence in support of this condition by demonstrating that predicted patenting growth is not correlated with pre-existing local trends in segregation. In Section IID we will run additional robustness checks to address remaining endogeneity concerns.

The right panels of Figure 6 show the reduced-form relationships between predicted patenting growth (1990–2010) and past changes in the three measures of

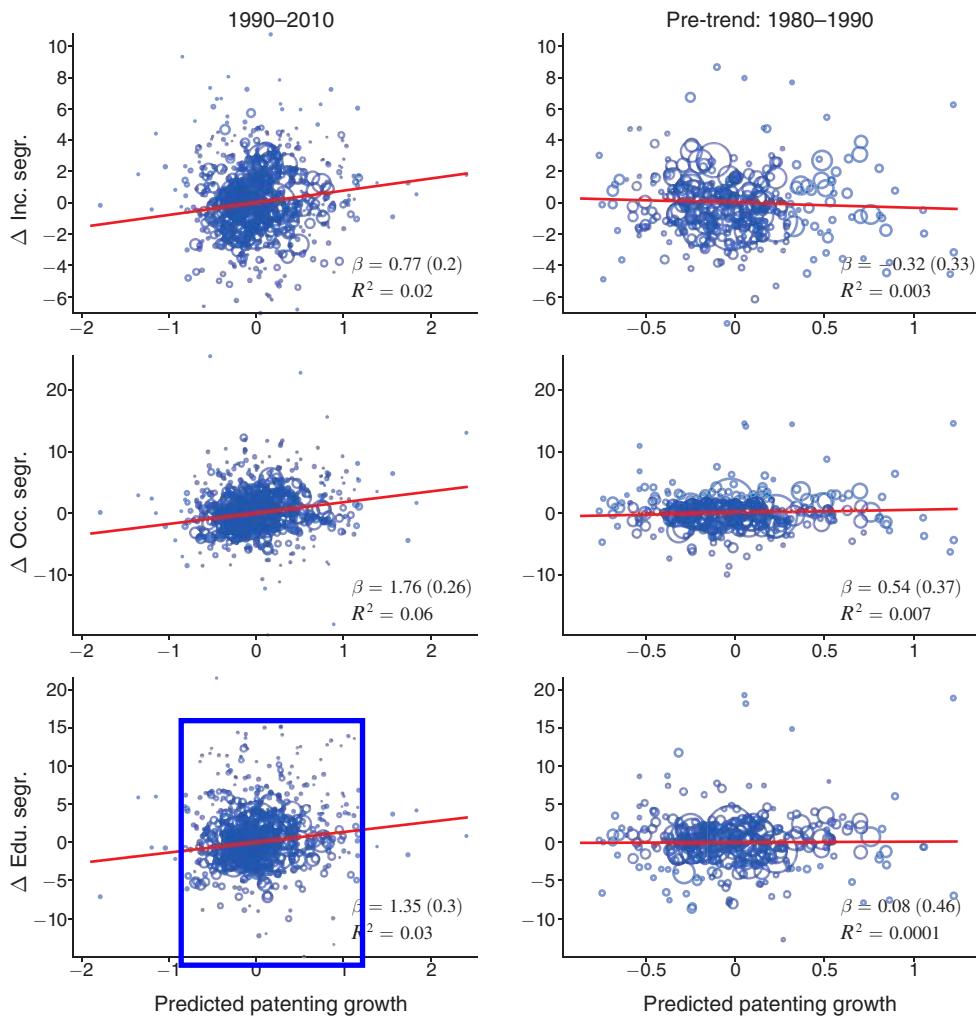


FIGURE 6. PREDICTED PATENTING GROWTH AND ECONOMIC SEGREGATION

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

economic segregation (1980–1990) once we partial out the full set of controls. For all three measures the slope of the regression line is statistically indistinguishable from zero. This suggests that innovation shocks in the years 1990–2010 are not correlated with the trajectory of economic segregation in the previous decade.

Online Appendix Table A.4 reports estimates of the corresponding OLS regressions in which we progressively introduce the set of controls. The estimated coefficients of predicted patenting growth are statistically indistinguishable from zero at the 5 percent level for all the specifications. Note that the years that 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, there are fewer observations in the pre-trend analysis. As we show in online Appendix Table A.5, our main results are robust to restricting the sample to cities for which 1980 data are available.

C. Instrumental Variable (IV) Results

We now use our instrument to explore the causal effects of innovation shocks on economic segregation in cities. The left panels of Figure 6 display the reduced-form relationships 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 the 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 percent 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 0.31 standard deviations of the over-time change in income segregation. The effect is even more pronounced for the measures of occupational (0.54 standard deviations) and educational (0.40 standard deviations) segregation.

The 2SLS estimates are larger than the ones in the OLS regressions. This fact suggests that unobservable factors simultaneously affecting 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 in the urban structure by tightening residential mortgages and 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 this variable is indeed informative of both the citation network in the early sample and the factors that control the evolution of segregation between 1990 and 2010, as discussed in Section IIB. 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 the coefficient stability test developed by Oster (2019) on the reduced-form regressions (shown in the left panels of Figure 6). Intuitively, this test proposes that when the inclusion of controls generates

TABLE 2—PATENTING GROWTH AND ECONOMIC SEGREGATION: IV RESULTS

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. $\Delta IncSegr$</i>						
Patenting growth	3.22 (0.59)	1.92 (1.00)	1.77 (0.90)	1.25 (0.60)	1.34 (0.59)	1.30 (0.61)
Share of college graduates		10.03 (5.56)	0.93 (6.28)	3.14 (4.82)	-6.18 (5.67)	-5.91 (5.72)
log CTs			0.60 (0.19)	-1.57 (1.02)	-1.65 (1.08)	-1.66 (1.08)
log households				2.09 (1.02)	1.79 (1.15)	1.80 (1.15)
log average income					4.90 (1.63)	4.87 (1.62)
Import exposure						0.04 (0.10)
<i>Panel B. $\Delta OccSegr$</i>						
Patenting growth	4.18 (0.53)	2.89 (0.77)	2.65 (0.66)	2.81 (0.65)	2.90 (0.62)	2.96 (0.61)
Share of college graduates		9.93 (4.61)	-4.47 (3.72)	-5.13 (3.87)	-14.57 (4.74)	-15.04 (4.98)
log CTs			0.95 (0.12)	1.61 (0.90)	1.53 (0.95)	1.55 (0.95)
log households				-0.63 (0.86)	-0.94 (0.90)	-0.96 (0.90)
log average income					4.96 (2.23)	5.00 (2.26)
Import exposure						-0.07 (0.13)
<i>Panel C. $\Delta EduSegr$</i>						
Patenting growth	3.62 (0.59)	2.50 (0.90)	2.37 (0.83)	2.24 (0.78)	2.41 (0.70)	2.27 (0.72)
Share of college graduates		8.62 (4.83)	0.35 (4.97)	0.89 (5.06)	-15.85 (5.48)	-14.71 (5.67)
log CTs			0.55 (0.22)	0.03 (1.35)	-0.11 (1.38)	-0.15 (1.40)
log households				0.50 (1.31)	-0.05 (1.29)	0.00 (1.32)
log average income					8.79 (2.47)	8.67 (2.42)
Import exposure						0.18 (0.15)
Observations	722	722	722	722	722	722
<i>First-stage estimates</i>						
Predicted patenting growth	0.71 (0.10)	0.64 (0.10)	0.65 (0.11)	0.59 (0.10)	0.59 (0.10)	0.59 (0.10)
F-stat	54.9	37.5	35.4	36.1	34.0	34.7
R ²	0.39	0.39	0.40	0.41	0.41	0.41

Notes: Regressions are weighted by total number of households in 1990. Controls are at 1990 values with the exception of import exposure (provided by Autor, Dorn, and Hanson 2013b). Standard errors clustered at the state level are in parentheses.

a significant drop in the estimated coefficient of interest and a small improvement 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 to bring the coefficient of interest to zero. All values are greater than the critical value of one, indicating that selection on unobservables would need to be greater than selection on observables to generate a null effect.²⁰

D. 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 2 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 geographic areas linked in the knowledge network have characteristics such as a similar industry structure, geographic proximity, common regulation, or exposure to other shocks that make it hard to disentangle the genuine effect of knowledge shocks from the influence of other factors that simultaneously affect 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 $\Lambda_c^{1990} = \{\lambda_{c,1}^{1990}, \dots, \lambda_{c,S}^{1990}\}$, where $\lambda_{c,\nu}^{1990}$ denotes the share of patents produced in c that belong to technology class ν . We then compute the 1990–2010 growth rate $g_{-c,\nu}$ of the number of grants in ν , considering only patents produced outside of c . Our control variable is then computed as

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

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.²¹ 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, information technology, etc.)

²⁰The 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).

²¹Hornbeck and Moretti (2018) use a similar variable as an instrument for local productivity shocks.

TABLE 3—PATENTING GROWTH AND ECONOMIC SEGREGATION: ROBUSTNESS

	$\Delta IncSegr$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patenting growth	1.30 (0.61)	1.52 (0.76)	1.13 (0.57)	1.62 (0.67)	1.74 (0.65)	1.35 (0.62)	1.22 (0.62)	2.02 (0.67)
Patenting growth (pre-trend)		-0.48 (0.47)						
Bartik-like variable			0.16 (0.60)					
Class shares controls (1990)	×	×	×	✓	×	×	×	×
Empl. shares controls (1990)	×	×	×	×	✓	×	×	×
Constr. instr. (No same class)	×	×	×	×	×	✓	×	×
Constr. instr. (No neighbor CZs)	×	×	×	×	×	×	✓	×
State fixed effects	×	×	×	×	×	×	×	✓
Baseline controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	722	722	706	706	722	722	722	722
<i>First-stage estimates</i>								
Predicted patenting growth	0.59 (0.10)	0.52 (0.09)	0.74 (0.16)	0.65 (0.13)	0.54 (0.10)	0.41 (0.10)	0.40 (0.09)	0.53 (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 the total number of households in 1990. Controls are at 1990 values, with the exception of import exposure (provided by Autor, Dorn, and Hanson 2013b). Missing observations in columns 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 are in parentheses.

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 the results are robust when controlling directly for the local share of patents in each of the eight 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 estimates 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 $d_{(b,\mu) \rightarrow (c,\nu)}^\tau = 0$ 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 geographic clustering of innovation that is reflected in the citation network. In column 7, we show that results are similar when using a version of the instrument that excludes citations from neighboring cities.²²

²²The fact that the estimates in columns 6 and 7 are close to the baseline results in column 1 also suggests that potential feedback effects between origin city-classes and destination city-classes due to loops in the citation net-

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 of innovation activities only through within-state variation. The estimated coefficient, reported in column 8, is positive, significant, and greater in magnitude.

E. 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 identified in this empirical framework, it is possible to estimate an upper bound for the inequality channel's contribution to the overall effect. Specifically, the extent to which the inequality channel is 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 income dispersion across households (measured by an increase in the city-level Gini index) owing, for example, to a jump in the college premium is reflected less than one-to-one in an increase in measured segregation. In other words, the change in income inequality provides an upper bound for the inequality channel's contribution to the change in income segregation.

To gain intuition into this claim, consider the effect of an increase in income inequality on income segregation when a city is perfectly segregated (i.e., 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, cities in the sample are generally imperfectly segregated. In this case, in the absence of any change in sorting, positive changes in inequality transmit less than one-to-one to changes in segregation. In the extreme case of a city in which the income distribution in each neighborhood is the same (which implies that income segregation is zero), an increase in income inequality would have no impact on income segregation.²³

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 the dependent variables. The impact of innovation on inequality (column 2) is less than

work are unlikely to be driving our results. In fact, these loops are more likely to emerge between combinations of city-classes that are close either technologically or geographically.

²³Note that in reality we observe changes in income segregation that are greater than contemporaneous changes in income inequality. This is because residents generally 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 under the condition that people are not allowed to relocate.

TABLE 4—INCOME SEGREGATION, INEQUALITY, AND PREDICTED PATENTING GROWTH

	$\Delta IncSegr$ (1)	$\Delta IncIneq$ (2)	$\Delta OccPrem$ (3)
Patenting growth	1.30 (0.60)	0.91 (0.40)	0.19* (0.06)
Baseline controls	✓	✓	✓
Estimation	2SLS	2SLS	2SLS
Observations	722	722	722

Notes: Regressions are weighted by the total number of households in 1990. Controls are at 1990 values, with the exception of import exposure (provided by Autor, Dorn, and Hanson 2013b). $\Delta IncSegr$ and $\Delta IncIneq$ are both computed using household income data and are expressed in terms of Gini points. $\Delta OccPrem$ is defined as the 1990–2010 change in the ratio between average earnings by workers in knowledge-intensive occupations and average earnings of workers in residual occupations (averages are computed from individual-level data from IPUMS). Standard errors clustered at the state level are in parentheses.

the effect on income segregation (column 1) by roughly one-third, suggesting that the sorting channel explains at least one-third, and possibly more, of the response of income segregation to local innovation shocks.²⁴ This is because 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 30 percent 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 by workers in knowledge-intensive occupations relative to workers in residual occupations in the local labor market, suggesting that the occupation premium is one of the channels that drives inequality up in response to innovation shocks.

III. 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 US cities between 1990 and 2010. In this section, we provide suggestive evidence for an economic mechanism that can drive this causal relationship. In particular, we show that local innovation shocks lead to a geographical clustering of knowledge-intensive jobs, inducing residents employed in knowledge-intensive occupations to relocate to reduce their commuting. The resulting increase in economic segregation is amplified by a localized response of rental prices and consumption amenities. This evidence provides guidance on what policies—such as expansion of transit infrastructure, loosening zoning restrictions, or public investment in local amenities—might be effective at mitigating the impact of knowledge-intensive activities on local spatial inequalities.

²⁴Note that income inequality and income segregation are both computed using the same household income data and are expressed in terms of Gini points; hence, the coefficients in Column 1 and 2 are directly comparable.

In line with the extensive body of literature on the centrality of localized idea flows for innovation (e.g., Jaffe, Trajtenberg, and Henderson 1993, Carlino and Kerr 2015), we interpret local innovation shocks as an increase in the returns from localized knowledge spillovers, which drives up the incentives of knowledge-intensive firms to cluster in space. Leveraging data from the NETS, we start by showing that—consistent with this interpretation—cities with higher innovation shocks display a more pronounced clustering of knowledge-intensive jobs in neighborhoods characterized by a high initial density of knowledge workers. We then demonstrate that this clustering of jobs coincides with a change in the residential choices of knowledge workers, who relocate toward neighborhoods with a lower commuting distance from these 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 residential clustering coincides with changes in both rental prices and measures of local consumption amenities (e.g., restaurants, food shops, and fitness centers). If different categories of residents have heterogeneous valuations of those amenities or a heterogeneous sensitivity of their residential choices to changes in rent (as estimated by Diamond 2016 and Almagro and Domínguez-Iino 2021, among others), these responses can significantly amplify the effect of innovation shocks on economic segregation.²⁵

A. 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 IIB identifies shocks to local patenting that result from the greater availability of ideas used as inputs to innovation. As proposed and extensively tested in the innovation literature, geographic proximity is a key determinant of knowledge exchange. Therefore, when new ideas become available and give rise to innovation opportunities, the incentives of knowledge-intensive firms to collocate increase and, for the marginal firm, outweigh the costs of colocation. This should result in a more pronounced clustering of knowledge-intensive jobs in neighborhoods that provide better opportunities for knowledge exchange.

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

$$(8) \quad \Lambda_{n,c,1990} = \frac{W_{n,c,1990}}{L_{n,c}},$$

²⁵In ongoing work (Berkes and Gaetani 2021) we develop an extension of the model in Ahlfeldt et al. (2015) 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 us to quantify the amplifying effects of endogenous rental prices and supply of residential amenities.

where $W_{n,c,1990}$ is the number of workers in knowledge-intensive occupations in 1990 and $L_{n,c}$ is the total amount of land (in square kilometers). This formulation mirrors the functional form for productivity externalities in Ahlfeldt et al. (2015) and Tsivanidis (2018), who model agglomeration forces—which increase neighborhood productivity—as a geometric function of employment density.²⁶

We then test the hypothesis that in cities that receive higher innovation shocks, knowledge-intensive jobs cluster in neighborhoods with stronger initial learning opportunities. Letting $\Delta s_{n,c}^w$ be the 1990–2010 change in the share of workers employed in knowledge-intensive occupations in neighborhood n of city c , we estimate the following equation:

$$(9) \quad \Delta s_{n,c}^w = \alpha_c^w + \beta^w \log(\Lambda_{n,c,1990}) + \gamma^w \log(\Lambda_{n,c,1990}) \\ \times \Delta \log(1 + Patents_c) + \epsilon_{n,c}^w,$$

where α_c^w is a city fixed effect. A positive sign for the coefficient of the interaction, γ^w , would suggest that neighborhoods with a higher initial density of knowledge-intensive jobs in cities with stronger innovation shocks have experienced a more pronounced shift toward knowledge-intensive occupations.

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 terms of magnitude. Combining the estimates of β^w and γ^w in column 2, we obtain that in cities at the ninetieth percentile of the distribution of innovation shocks, an increase of one within-city standard deviation in the initial density of knowledge workers is associated with a 1.01 percentage point shift toward knowledge-intensive occupations, equal to 0.19 within-city standard deviations of $\Delta s_{n,c}^w$. The corresponding effect in cities at the fifth percentile of the distribution of innovation shocks is significantly lower (0.35 percentage points, equal to 0.07 within-city standard deviations of $\Delta s_{n,c}^w$).

These results are consistent with a theoretical framework in which neighborhood-level learning opportunities ($\Lambda_{n,c,1990}$) 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. This complementarity implies that stronger innovation shocks increase the productivity of knowledge workers in neighborhoods with higher learning opportunities relative to neighborhoods with lower opportunities. This induces the marginal knowledge worker to relocate to a neighborhood where agglomeration forces are stronger. At the same time, this relocation results in an increase in the rental price of office space. As a consequence, the marginal worker in a residual occupation

²⁶ 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 (i.e., a city 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 such that terms from surrounding neighborhoods are unlikely to contribute significantly to the measure in equation (8).

TABLE 5—CLUSTERING OF KNOWLEDGE WORKERS AND RESIDENTS

	Workers ($\Delta s_{n,c}^w$)		Residents ($\Delta s_{n,c,0-5}^r$)	
	(1)	(2)	(3)	(4)
$\log(\Lambda_{n,c,1990})$	0.30 (0.03)	0.28 (0.03)	0.26 (0.13)	-0.004 (0.12)
$\log(\Lambda_{n,c,1990}) \times \Delta \log(1 + Patents_c)$	0.11 (0.06)	0.23 (0.07)	0.82 (0.20)	1.90 (0.26)
CZ fixed effects	✓	✓	✓	✓
Estimation	OLS	2SLS	OLS	2SLS
Observations	60,452	60,452	59,666	59,666
Clusters	714	714	714	714
R^2	0.05		0.18	
First-stage F -stat		136.85		119.7

Notes: $\Delta s_{n,c}^w$ is defined as the 1990–2010 change in the percentage of workers employed in knowledge-intensive occupations in census tract n . $\Delta s_{n,c,0-5}^r$ is defined as the change between 1990 and 2010 in the percentage of residents employed in knowledge-intensive occupations in neighborhoods at less than five minutes of commuting distance from n . 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 are in parentheses.

whose productivity is unaffected by the innovation shock is induced to relocate to a neighborhood where learning opportunities are weaker. Overall, these forces lead to a more pronounced increase in the share of knowledge workers ($s_{n,c}^w$) in neighborhoods with higher learning opportunities in cities with stronger innovation shocks.

B. Response in the Residential Choices of Knowledge Workers

The estimates of equation (9) suggest that innovation shocks are associated with 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 in 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, consistent with this intuition, stronger innovation shocks induce knowledge workers to relocate their residences 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.

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 five minutes of commuting distance from neighborhood n as follows:

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

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, γ^r , implies that in cities with stronger innovation shocks residents employed in knowledge-intensive occupations relocate more heavily to neighborhoods with a short (zero- to five-minute) commuting distance from locations with high initial density of knowledge workers (captured by $\Lambda_{n,c,1990}$) 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 1990.²⁷ The coefficient of the interaction term is estimated to 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 ninetieth percentile of the distribution of innovation shocks, an increase of one within-city standard deviation in the initial density of knowledge workers is associated with a 3.33 percentage point increase in the share of residents in knowledge-intensive occupations, equal to 0.68 within-city standard deviations of $\Delta s_{n,c,0-5}^r$. The corresponding effect in cities at the fifth percentile of the distribution of innovation shocks is opposite in sign (−2.05 percentage points, equal to a decrease of 0.42 within-city standard deviations of $\Delta s_{n,c,0-5}^r$).

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, which we will 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 distances 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 γ^r and 95 percent confidence intervals. The figure indicates 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 distance, losing statistical significance for neighborhoods located 15–20 minutes away and turning negative and significant for neighborhoods located more than 25 minutes away.

C. Effect on Rental Price of Housing and Consumption Amenities

The estimates obtained from equation (10) suggest a significant change in the geographic segmentation of residents by occupation groups that can partially explain the effect of innovation shocks on economic segregation. As previously discussed,

²⁷In this case we also multiply the coefficients by 100 to interpret them as percentage points.

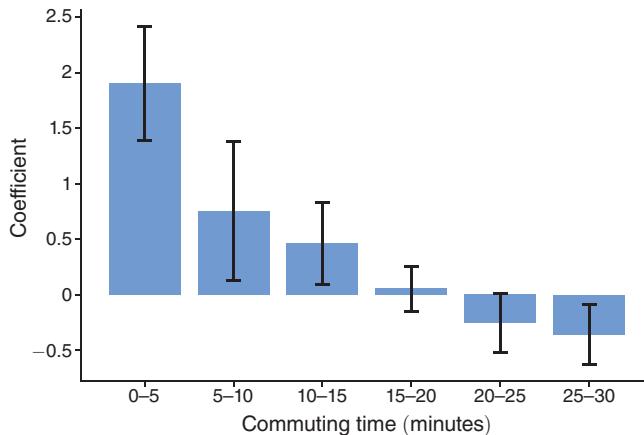


FIGURE 7. CLUSTERING OF RESIDENTS BY COMMUTING DISTANCE

Notes: The plot shows the 2SLS point estimates and 95 percent confidence intervals of the coefficient of the interaction of $\log(\Lambda_{n,c,1990})$ and $\Delta\log(1 + \text{Patents}_c)$ in specifications analogous to equation (10). The dependent variable is defined as $\Delta s_{n,c,m_1-m_2}^r$, where $(m_1 - m_2)$ is the commuting distance bin indicated on the horizontal axis. Observations are weighted by the total number of residents in 1990. Standard errors are clustered at the city level.

the effect is more pronounced for the clustering of residents than for the clustering of workers. In this subsection we provide suggestive evidence for an amplification channel that might account for this difference. We show that the relocation of knowledge residents coincides with a response in rental prices as well as measures of local consumption amenities. This response can work as a powerful amplification channel for economic segregation if, as proposed and estimated in the literature (Diamond 2016; Couture et al. 2020; 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 rental prices.

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 as the dependent variable:

$$(11) \quad \Delta q_{n,c,0-5} = \alpha_c^q + \beta^q \log(\Lambda_{n,c,1990}) + \gamma^q \log(\Lambda_{n,c,1990}) \\ \times \Delta\log(1 + \text{Patents}_c) + \epsilon_{n,c}^q,$$

where $\Delta q_{n,c,0-5}$ is the log change in the average rental price of housing in neighborhoods at less than five minutes of commuting distance from n .

Results are reported in columns 1 and 2 of Table 6. The effect is positive and statistically significant, with the point estimate in the IV regression implying that in cities at the ninetieth percentile of the distribution of innovation shocks, a increase of one within-city standard deviation in the 1990 density of knowledge workers is associated with a 10.6 percent increase in the rental price of housing, equivalent to 0.52 within-city standard deviations. The corresponding effect in cities at the fifth

TABLE 6—RESPONSE OF RENTAL PRICE OF HOUSING AND CONSUMPTION AMENITIES

	Rental price ($\Delta q_{n,c,0-5}$)		Restaurants ($\Delta A_{n,c,0-5}^i$)		Food shops ($\Delta A_{n,c,0-5}^i$)		Fitness centers ($\Delta A_{n,c,0-5}^i$)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\log(\Lambda_{n,c,1990})$	0.012 (0.007)	0.001 (0.005)	0.207 (0.027)	0.177 (0.022)	0.023 (0.006)	0.021 (0.005)	0.063 (0.018)	0.051 (0.012)
$\log(\Lambda_{n,c,1990}) \times \Delta \log(1 + Patents_c)$	0.015 (0.007)	0.059 (0.014)	0.069 (0.040)	0.193 (0.049)	0.018 (0.010)	0.026 (0.011)	0.032 (0.018)	0.083 (0.039)
CZ fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Estimation	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Observations	59,469	59,469	59,667	59,667	59,667	59,667	59,667	59,667
Clusters	714	714	714	714	714	714	714	714
R ²	0.24		0.23		0.15		0.14	
First-stage F-stat		119.7		119.7		119.7		119.7

Notes: $\Delta q_{n,c,0-5}$ is defined as the 1990–2010 log change in the average rental price of housing in neighborhoods at less than five minutes of commuting distance from n . $\Delta A_{n,c,0-5}^i$ is defined as in equation (12), where i refers to “restaurants” (columns 3 and 4), “food shops” (columns 5 and 6), and “fitness centers” (columns 7 and 8). Observations are weighted by the total number of residents in 1990. Standard errors clustered at the CZ level are in parentheses.

percentile of the distribution of innovation shocks is opposite in sign (−6.0 percent, equal to a decrease of 0.29 within-city standard deviations).

The upper-left panel of Figure 8 reports the estimated coefficients of the interaction term (γ^q) for the average rental prices located at various bins of commuting distance. The results are strongly aligned with those presented in Figure 7, with the coefficient of the interaction term steeply declining with distance and becoming negative and significant at locations requiring commutes longer than 25 minutes.

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 interaction of a fixed supply and an increasing demand for residential and office space that drives rents up. A complementary view is that knowledge workers have a higher opportunity cost of time (Su 2022), which increases their willingness to pay for housing with lower commuting distances from the neighborhoods in which more knowledge-intensive jobs are available. An alternative 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., Diamond 2016; Baum-Snow and Hartley 2020; Couture and Handbury 2020) or because of nonhomothetic demand (e.g., Couture et al. 2020), their appearance can increase the desirability of locations and drive up the rental price of housing.

To provide suggestive evidence consistent with the existence of this amplification channel, we estimate specifications analogous to equation (11) with the dependent variable defined as

$$(12) \quad \Delta A_{n,c,0-5}^i = 1,000 \times \frac{\Delta K_{n,c,0-5}^i}{R_{n,c,0-5,1990}},$$

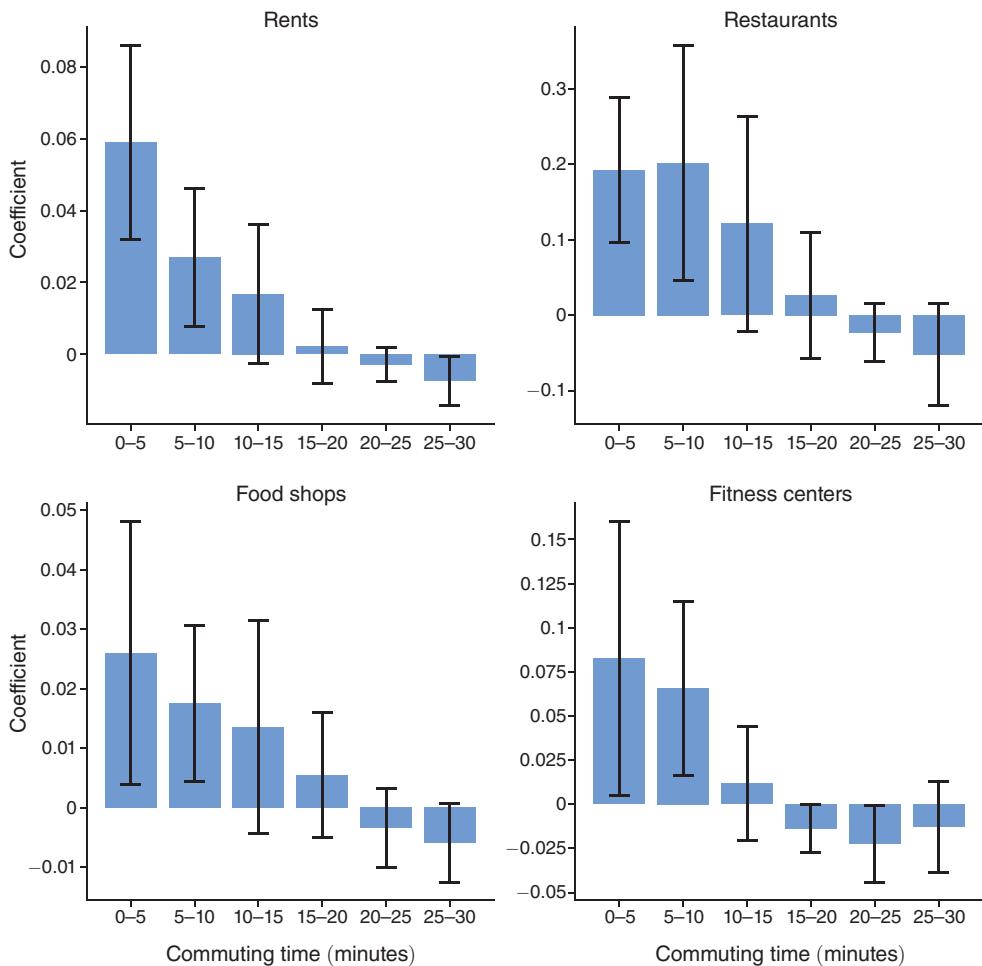


FIGURE 8. RENTAL PRICE OF HOUSING AND CONSUMPTION AMENITIES BY COMMUTING DISTANCE

Notes: The plot shows 2SLS point estimates and 95 percent confidence intervals of the coefficient of the interaction of $\log(\Lambda_{n,c,1990})$ and $\Delta \log(1 + \text{Patents}_c)$ in specifications analogous to equation (11). The dependent variable is defined as the percentage change in average rent (upper-left panel), and $\Delta A_{n,c,m_1-m_2}^i$, where i refers to the following categories—“restaurants” (upper-right panel), “food shops” (lower-left panel), and “fitness centers” (lower-right panel)—in neighborhoods located 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.

where $\Delta K_{n,c,0-5}^i$ is the change between 1990 and 2010 in the number of establishments in category i located in neighborhoods at commuting distance between zero and five minutes from n , and $R_{n,c,0-5,1990}$ 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 online Appendix B.2 for details regarding the definitions of these categories).

Regression results are reported in columns 3–8 of Table 6. In the IV regressions (columns 4, 6, and 8), the coefficients of the interaction range from 0.026 per 1,000 residents for food shops to 0.193 per 1,000 residents for restaurants. For comparison,

the average within-city standard deviations in $\Delta A_{n,c,0-5}^i$ are equal to 1.36, 0.37, and 0.46 per 1,000 residents for restaurants, food shops, and fitness centers, respectively.

The upper-right and bottom panels of Figure 8 plot the estimated coefficients of the interaction term at different bins of commuting distance. The response in the availability of consumption amenities displays a spatial decay analogous to the one observed for the rental price of housing. This suggests that innovation shocks are a meaningful source of variation in the evolution of consumption amenities in cities and that the endogenous arrival of local amenities can work as an amplification channel for the effect of innovation shocks on economic segregation.

IV. Conclusion

In recent decades, policy makers have been increasingly promoting the development of local innovation-based activities by, for example, supporting aggressive bids to attract knowledge-intensive firms (Slattery and Zidar 2020). In many cases local communities have responded to these 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 in urban areas. Our IV estimates suggest that in the cross-section of US 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 0.31 standard deviations of the over-time change in income segregation. We provide suggestive evidence of a potential mechanism underlying this relationship: the expansion of local innovation possibilities increases the incentives of knowledge-intensive firms to spatially cluster to take advantage of localized learning opportunities. Because of commuting considerations, this clustering of employment increases residential sorting, with the effect amplified by the endogenous response of rental prices 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 geographic divide between individuals with different levels of income has worsened, possibly exacerbating differences in terms of access to education, health, and consumption amenities. This contributes 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, or a change in industry, occupation, or labor force status) are important avenues 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 simultaneously increase the accessibility of more peripheral areas. Second, loosening restrictive zoning policies could dampen the sharp response

in rental housing prices, diminishing the impact on residential sorting. Third, the amplification generated by the inflow of local consumption amenities could be weakened by redistributive policies that limit divergences in the quality of amenities across neighborhoods.

The effectiveness of these policies should be evaluated within a framework that considers general equilibrium interactions. In ongoing work (Berkes and Gaetani 2021), we are developing 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 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—which is essential to the economic prosperity of countries and regions—are distributed in a more sustainable and inclusive manner.

REFERENCES

- Acemoglu, Daron, Ufuk Akgit, and William R. Kerr.** 2016. “Innovation Network.” *Proceedings of the National Academy of Sciences* 113 (41): 11483–88.
- Acevedo-Garcia, Dolores, Kimberly A. Lochner, Theresa L. Osypuk, and S.V. Subramanian.** 2003. “Future Directions in Residential Segregation and Health Research: A Multilevel Approach.” *American Journal of Public Health* 93 (2): 215–21.
- Aghion, Philippe, Ufuk Akgit, Antonin Bergeaud, Richard Blundell, and David Hemous.** 2019. “Innovation and Top Income Inequality.” *Review of Economic Studies* 86 (1): 1–45.
- Agrawal, Ajay, Alberto Galasso, and Alexander Oettl.** 2017. “Roads and Innovation.” *Review of Economics and Statistics* 99 (3): 417–34.
- Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf.** 2015. “The Economics of Density: Evidence from the Berlin Wall.” *Econometrica* 83 (6): 2127–89.
- Alexander, Diane, and Janet Currie.** 2017. “Is It Who You Are or Where You Live? Residential Segregation and Racial Gaps in Childhood Asthma.” *Journal of Health Economics* 55: 186–200.
- Almagro, Milena, and Tomás Domínguez-Iino.** 2021. “Location Sorting and Endogenous Amenities: Evidence from Amsterdam.” Unpublished.
- Autor, David.** 2019. “Work of the Past, Work of the Future.” NBER Working Paper 25588.
- Autor, David, David Dorn, and Gordon H. Hanson.** 2013a. “The China Syndrome: Local Labor Market Effects of Import Competition in the United States.” *American Economic Review* 103 (6): 2121–68.
- Autor, David, David Dorn, and Gordon H. Hanson.** 2013b. “The China Syndrome: Local Labor Market Effects of Import Competition in the United States: Dataset.” *American Economic Review*. <https://www.aeaweb.org/articles?id=10.1257/aer.103.6.2121>.
- Baum-Snow, Nathaniel.** 2007. “Did Highways Cause Suburbanization?” *Quarterly Journal of Economics* 122 (2): 775–805.
- Baum-Snow, Nathaniel, and Daniel Hartley.** 2020. “Accounting for Central Neighborhood Change, 1980–2010.” *Journal of Urban Economics* 117: 103228.
- Baum-Snow, Nathaniel, and Byron F. Lutz.** 2011. “School Desegregation, School Choice, and Changes in Residential Location Patterns by Race.” *American Economic Review* 101 (7): 3019–46.
- Baum-Snow, Nathaniel, and Ronni Pavan.** 2013. “Inequality and City Size.” *Review of Economics and Statistics* 95 (5): 1535–48.
- Berkes, Enrico, and Ruben Gaetani.** 2021. “A Quantitative Model of Income Segregation in a Knowledge Economy.” Unpublished.
- Berkes, Enrico, and Ruben Gaetani** 2023. “Replication Data for: Income Segregation and the Rise of the Knowledge Economy.” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E152903V1>.

- Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb.** 2020. "Are Ideas Getting Harder to Find?" *American Economic Review* 110 (4): 1104–44.
- Borusyak, Kirill, Peter Hull, and Xavier Jaravel.** 2018. "Quasi-experimental Shift-Share Research Designs." NBER Working Paper 24997.
- Bureau of Economic Analysis.** 2018. "Value Added by Industry as a Percentage of Gross Domestic Product." Bureau of Economic Analysis. https://www.bea.gov/sites/default/files/2018-04/GDPby-Ind_VA_1947-2017.xlsx (accessed January 21, 2022).
- Carlino, Gerald, and William R. Kerr.** 2015. "Agglomeration and Innovation." In *Handbook of Regional and Urban Economics*, Vol. 5, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, 349–404. Amsterdam: Elsevier.
- Catalini, Christian.** 2018. "Microgeography and the Direction of Inventive Activity." *Management Science* 64 (9): 4348–64.
- Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz.** 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106 (4): 855–902.
- Coile, Courtney C., and Mark G. Duggan.** 2019. "When Labor's Lost: Health, Family Life, Incarceration, and Education in a Time of Declining Economic Opportunity for Low-Skilled Men." *Journal of Economic Perspectives* 33 (2): 191–210.
- Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst.** 2020. "Income Growth and the Distributional Effects of Urban Spatial Sorting." NBER Working Paper 26142.
- Couture, Victor, and Jessie Handbury.** 2020. "Urban Revival in America." *Journal of Urban Economics* 119: 103267.
- Davis, Donald R., Jonathan I. Dingel, Joan Monras, and Eduardo Morales.** 2019. "How Segregated is Urban Consumption?" *Journal of Political Economy* 127 (4): 1684–1738.
- Diamond, Rebecca.** 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000." *American Economic Review* 106 (3): 479–524.
- Dingel, Jonathan I., and Felix Tintelnot.** 2020. "Spatial Economics for Granular Settings." NBER Working Paper 27287.
- Duranton, Gilles, and Matthew A. Turner.** 2012. "Urban Growth and Transportation." *Review of Economic Studies* 79 (4): 1407–40.
- Eckert, Fabian, Sharat Ganapati, and Conor Walsh.** 2020. "Skilled Scalable Services: The New Urban Bias in Economic Growth." Unpublished.
- Florida, Richard L.** 2002. *The Rise of the Creative Class*. New York, NY: Basic Books.
- Florida, Richard L., and Charlotta Mellander.** 2015. *Segregated City: The Geography of Economic Segregation in America's Metros*. Toronto: Martin Prosperity Institute.
- Fogli, Alessandra, and Veronica Guerrieri.** 2019. "The End of the American Dream? Inequality and Segregation in US Cities." NBER Working Paper 26143.
- Gans, Joshua, and Andrew Leigh.** 2019. *Innovation+ Equality: How to Create a Future that is More Star Trek Than Terminator*. Cambridge, MA: MIT Press.
- Gaubert, Cecile, Patrick M. Kline, Damián Vergara, and Danny Yagan.** 2021. "Trends in US Spatial Inequality: Concentrating Affluence and a Democratization of Poverty." NBER Working Paper 28385.
- Glaeser, Edward L., and Joshua D. Gottlieb.** 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4): 983–1028.
- Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman, and Andrei Shleifer.** 1992. "Growth in Cities." *Journal of Political Economy* 100 (6): 1126–52.
- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift.** 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110 (8): 2586–2624.
- Gorback, Caitlin.** 2020. "Your Uber Has Arrived: Ridesharing and the Redistribution of Economic Activity." Unpublished.
- Hornbeck, Richard, and Enrico Moretti.** 2018. "Who Benefits from Productivity Growth? Direct and Indirect Effects of Local TFP Growth on Wages, Rents, and Inequality." NBER Working Paper 24661.
- Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson.** 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (3): 577–98.
- Jargowsky, Paul A.** 1996. "Take the Money and Run: Economic Segregation in US Metropolitan Areas." *American Sociological Review* 61 (6): 984–98.

- Jones, Benjamin F.** 2009. "The Burden of Knowledge and the 'Death of the Renaissance Man': Is Innovation Getting Harder?" *Review of Economic Studies* 76 (1): 283–317.
- Jones, Charles I.** 1995. "R and D-Based Models of Economic Growth." *Journal of Political Economy* 103 (4): 759–84.
- Jones, Charles I., and Jihee Kim.** 2018. "A Schumpeterian Model of Top Income Inequality." *Journal of Political Economy* 126 (5): 1785–1826.
- Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman.** 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment." *Quarterly Journal of Economics* 116 (2): 607–54.
- Luxen, Dennis, and Christian Vetter.** 2011. "Real-Time Routing with OpenStreetMap Data." In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, edited by Divyakant Agrawal, Isabel Cruz, Christian S. Jensen, Eyal Ofek, and Egeman Tanin, 513–16. New York, NY: ACM.
- Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles.** 2020. "IPUMS National Historical Geographic Information System: Version 15.0." IPUMS. <http://doi.org/10.18128/D050.V15.0> (accessed August 2020).
- Moretti, Enrico.** 2012. *The New Geography of Jobs*. Boston, MA: Houghton Mifflin Harcourt.
- Oster, Emily.** 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics* 37 (2): 187–204.
- Powell, Walter W., and Kaisa Snellman.** 2004. "The Knowledge Economy." *Annual Review of Sociology* 30: 199–220.
- Reardon, Sean F., and Kendra Bischoff.** 2016. "The Continuing Increase in Income Segregation, 2007–2012." Stanford Center for Education Policy Analysis. <https://cepa.stanford.edu/content/continuing-increase-income-segregation-2007-2012> (accessed June 20, 2016).
- Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman.** 2019. "Cognitive Hubs and Spatial Redistribution." NBER Working Paper 26267.
- Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek.** 2021. "IPUMS USA: Version 11.0." IPUMS. <https://doi.org/10.18128/D010.V11.0> (accessed September 2020).
- Slattery, Caitlin, and Owen Zidar.** 2020. "Evaluating State and Local Business Incentives." *Journal of Economic Perspectives* 34 (2): 90–118.
- Su, Yichen.** 2022. "The Rising Value of Time and the Origin of Urban Gentrification." *American Economic Journal: Economic Policy* 14 (1): 402–39.
- Tsivanidis, Nick.** 2018. "The Aggregate and Distributional Effects of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio." Unpublished.
- United States Patent and Trademark Office.** 2015. "Bulk Data Products." United States Patent and Trademark Office. <https://bulkdata.uspto.gov/> (accessed October 2015).
- Walls and Associates.** 2014. "National Establishment Time-Series (NETS) Database." Ewing Marion Kauffman Foundation. <https://maryannfeldman.web.unc.edu/data-sources/longitudinal-databases/national-establishment-time-series-nets/> (accessed March 2017).
- White, Michael J.** 1983. "The Measurement of Spatial Segregation." *American Journal of Sociology* 88 (5): 1008–18.