International Inequalities Institute

Superstar Cities and Left-Behind Places: Disruptive Innovation, Labor Demand, and Interregional Inequality

Tom Kemeny and Michael Storper

Working paper 41

February 2020

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

LSE International Inequalities Institute

The International Inequalities Institute (III) based at the London School of Economics and Political Science (LSE) aims to be the world's leading centre for interdisciplinary research on inequalities and create real impact through policy solutions that tackle the issue. The Institute provides a genuinely interdisciplinary forum unlike any other, bringing together expertise from across the School and drawing on the thinking of experts from every continent across the globe to produce high quality research and innovation in the field of inequalities.

In addition to our working papers series all these publications are available to download free from our website: www.lse.ac.uk/III

For further information on the work of the Institute, please contact the Institute Manager, Liza Ryan at e.ryan@lse.ac.uk

International Inequalities Institute

The London School of Economics and Political Science

Houghton Street

London

WC2A 2AE

Email: Inequalities.institute@lse.ac.uk

Web site: www.lse.ac.uk/III

@LSEInequalities

Abstract

After a long period of convergence, around 1980, inter-place gaps in economic well-being in the United States began to increase. This rising inequality offers a rich terrain to explore causality in regional economics and development theory. This paper presents new, long-run evidence on interregional inequality that highlights the need to situate the current moment in a context of episodic alternations between convergence and divergence. In light of this evidence, the paper revisits the theoretical literature, finding gaps in existing supply- and demand-side models. A demand-led perspective can be strengthened by integrating a primary role for disruptive technological change. We posit a theory of alternating waves, where major technology shocks initially concentrate, and eventually deconcentrate, demand for skilled workers performing complementary tasks. Labor supply responds to these centripetal and centrifugal forces. These reversals yield the observed patterns of rising and falling interregional inequality. We trace out the implications of this theory in both academic and policy terms.

Keywords: cities, income inequality, economic geography, regional development, convergence
JEL Codes: R11, R12, O33, N9

1. From convergence to divergence

After a long period of convergence, around 1980, inter-place gaps in economic well-being in the United States began to increase (Browne, 1989). This shift from convergence to divergence has been observed in other developed economies, including Sweden (Enflo and Roses, 2015), Spain (Martinez-Galarraga et. al, 2015), as well as the EU as a whole (Roses and Wolf, 2018), and certain developing countries, including Mexico (Aguilar-Retureta, 2016). Per capita incomes and average wages are the principal indicators used to describe these gaps; a growing literature shows that disparities manifest along other dimensions, including the distribution of and returns to skill (Autor, 2019), as well as intergenerational mobility (Chetty et al., 2014; Connor, 2018).

In popular and scholarly accounts, these patterns are commonly described in terms of a split between prosperous ‘superstar’ metropolises and left-behind places – deindustrialized, shrinking cities and struggling rural areas. In the U.S., a contrast is often drawn between coastal cities and the rest of the country, although this glosses over the emergence of prosperous regions between the coasts, including Austin, Denver, and Houston; in the UK, it is seen as an acceleration of the classical split between London and the north; in France, the reprise of the dominance of the Paris region. Particularities aside, these narratives share the idea that certain groups in left-behind areas are stagnating in terms of income and life chances, while highly-educated workers everywhere, but especially those in certain dense, urban agglomerations continue to outdistance the rest. Such narratives line up with academic studies that find that talent, youth, wealth and innovation today are flowing to a limited set of mostly large metropolitan areas (Moretti, 2013; Diamond, 2016; Florida, 2017). In the extreme versions of divergence, a small set of superstar cities are differentiating themselves from the rest of the countries in which they are embedded. A growing, vibrant conversation is underway about this polarization, it is no doubt additionally fueled by a growing sense that this inequality is linked to rising populism and political upheaval observed in a wide range of countries (Rodriguez-Pose, 2018; Lee et al, 2018; Broz et al., 2019).

The academic analysis of causal mechanisms is split into two primary narratives: one emphasizing labor supply, and the other labor demand. Those centered on labor supply explain rising inequality as a consequence of contemporary skilled workers' locational choices. The demand-side explanations argue that the divergence is caused by the geography of skill-biased technological change.

This paper makes two contributions to this debate. First, we present new, consistent long-run evidence on patterns of U.S. interregional inequality. The switch from convergence to divergence in 1980 is not an isolated historical experience. Rather, a previous phase of rising interregional inequality took shape during the mid-19th century to early 20th century. A second contribution lies in our reconstruction of theory in light of this evidence. While existing demand- and supply-side theories of interregional inequality highlight important mechanisms at work, neither offers an account that can accommodate the kinds of alternations we observe in the data. To address this problem, we link existing demand-led explanations of recent regional divergence, such as Baum-Snow et al (2018), with economic historians' account of the process of technological change and growth (Mokyr, 1990; Perez, 2010), and geographers' pioneering work on the dynamics of American industrial location (Pred, 1970; Storper and Walker, 1989). We argue that technology is the primary mechanism regulating episodic shifts between convergence and divergence. Divergence phases are driven by disruptive technology shocks, such as those around electrification that took place after the 1860s, and those centered on digital technologies refined over the last third of the twentieth century. Convergence phases follow later on, as formerly new technologies become routinized and codified, allowing them to diffuse throughout the economy. The two phases are characterized by different locational choices of skilled workers, which drive the geography of incomes, housing costs, and amenities. From a policy perspective, a crucial insight of this model is that fundamentally different dynamics are at play in different phases of the cycle. It is therefore crucial to determine where in the cycle we are, as each implies different possibilities and policy levers.

2. Eight dimensions of the switch between convergence and divergence in the USA

Using public-use Census microdata aggregated to the level of 1990-vintage Commuting Zones (CZs), in this section we highlight eight key dimensions of the evolution of interregional economic divergence.[1]

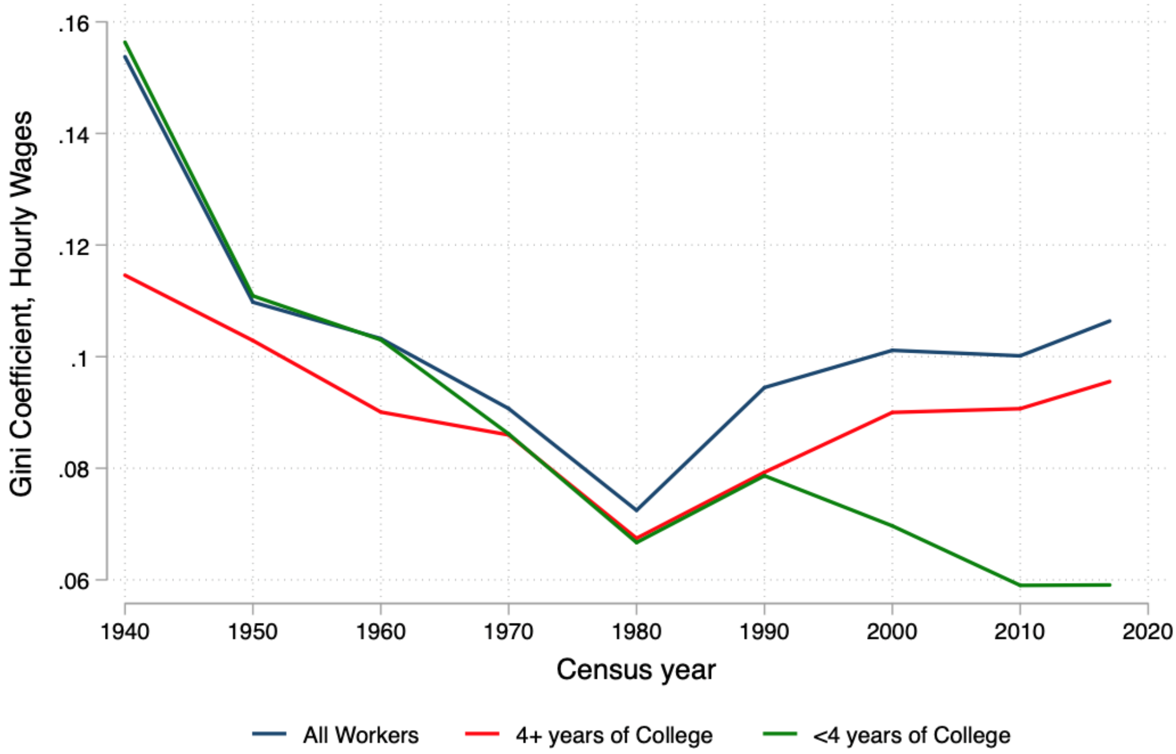
2.1 Nominal interregional income convergence switches to divergence around 1980

Figure 1 visualizes Gini coefficients of average hourly wage and salary income, tracing the evolution of interregional inequality between 1940 and 2017. Focusing on the blue line representing estimates for all in-sample workers, the figure shows that, from 1940 until 1980, regional differences in incomes were in decline. From 1980 forward, gaps begin to grow if we weight units by population (as we do in the figure); convergence still ends in 1980 if we do

[1] Commuting zones are determined according to the intensity of travel patterns and distinguished by weak inter-area commuting (Tolbert and Sizer, 1996). As compared with States, municipalities, or counties, the 722 CZs analyzed here may more closely delineate functionally linked economic regions. Unlike Metropolitan Areas, CZs also offer wider coverage; our set of locations cover the entirety of the contiguous continental U.S. Our underlying data are individual- and household-level public-use microdata obtained from successive Decennial Censuses, as well as the American Community Survey, courtesy of IPUMS (Ruggles et al., 2018). These data span the period 1940 to 2017. Details of the construction of our analytical sample are found in the Data Appendix.

not differentiate by population, replaced not by divergence but by stability. This distinction suggests that large metropolitan areas contribute significantly to the contemporary divergence.

Figure 1. Evolution of interregional income inequality (σ-convergence), overall and by education, 1940--2017



Note: N=722 Commuting Zones (CZs). Based on year-specific Gini coefficients estimated using average estimated hourly wage and salary income for all in-sample workers in 1990-vintage CZs, weighted by population. Incomes are adjusted for inflation to 2015 dollars using Bureau of Labor Statistics CPI. Source data are IPUMS public use extracts of Decennial Censuses and the American Community Survey. Further details in the text and in the Data Appendix.

At the scale of Census Regions, States, and cities, researchers began noting an interruption of post-war convergence soon after it began in the 1980s (i.e., Browne, 1989; Garnick, 1990; Drennan et al., 1996). Back then, most imagined it to be a brief aberration in a secular drive towards convergence (i.e. Carlino, 1992). Today we know that it has continued over at least 40 years. In the last decade, this realization has spurred a growing body of work, with contributions by economists, sociologists and geographers (i.e., Moretti, 2012, 2013; Kemeny and Storper, 2012; Diamond, 2016; Ganong and Shoag, 2017; Schwartzman, 2017; Giannone, 2017; Storper, 2018; Manduca, 2019; Autor, 2019). Less attention has been given to a consistent framework for understanding the causes of convergence in the 1940-80 period and how they might differ from mechanisms operating from 1980 onward.

2.2 Regional income performance in the 1980s involved turbulence and leapfrogging

We can qualify the basic picture shown in Figure 1 in several ways. One is to consider distributional dynamics. Figure 1 captures what Sala-i-Martin (1990) describes as σ-convergence, in that it tracks changes in overall dispersion. β-convergence is of additional interest, describing whether the pattern of growth is progressive, such that growth among lower-income economies exceeds growth in richer locations. The existence of β-convergence is necessary for σ-convergence (Sala-i-Martin, 1996). However, it is not sufficient: relatively larger growth among poor locations can result in their leapfrogging over richer locations, which is consistent with stable or growing overall dispersion. To unpack σ-convergence in our data, we adopt a method proposed by Jenkins and Van Kerm (2006), which additively decomposes changes in Gini coefficients to contributions from changes in β-convergence and positional mobility (or leapfrogging).[2] This accounting framework has the advantage of being non-parametric – unlike regression-based β-convergence estimates, it does not assume a linear growth process (O'Neill and Van Kerm, 2008).

Table 1. Decomposition of changes in interregional income inequality (σ-convergence) into contributions from leapfrogging and β-convergence, US Commuting Zones, 1940-1980

	Year Span	ΔGini (σ)	Leapfrogging	Progressivity (β)
Panel 1. Major Periods				
	1940-1980	-49.3%	5.7%	55.1%
	1980-2017	51.5%	30.9%	-20.6%
Panel 2. By Decade				
	1940-1950	-26.1%	2.5%	28.7%
	1950-1960	-1.6%	5.6%	7.2%
	1960-1970	-10.2%	4.2%	14.4%
	1970-1980	-20.4%	4.5%	24.9%
	1980-1990	32.5%	16.8%	-15.7%
	1990-2000	7.8%	2.8%	-5.0%
	2000-2010	0.6%	2.2%	1.6%
	2010-2017	6.7%	1.5%	-5.2%

Note: N=722 Commuting Zones. Values in the table are percentage changes from initial-period Gini coefficients. Gini coefficients are conventional in terms of weighting poorer units, i.e. estimated with ν=2. Values are calculated using the Stata program DSGINIDECO (Jenkins and Van Kerm, 2009). Actual Gini coefficients and bootstrapped standard errors are shown in Appendix B. For each Commuting Zone, 'income' is defined as estimated average hourly wage and salary income. Incomes are adjusted for inflation to 2015 dollars using Bureau of Labor Statistics CPI. Calculations are weighted by population in the initial period in question. Source data are public-use (IPUMS) extracts of the Decennial Censuses and American Community Survey. Data details are found in Appendix A

Table 1 reports results of this decomposition procedure. To support substantive interpretation, reported values are percentage changes from initial period Gini coefficients.[3] Panel 1 of the table highlights the two distinct phases in the 1940 to 2017 period. Between 1940 and 1980, regional disparities in average hourly wages decline by about half. Between

[2] For a derivation behind this approach, readers are directed to Jenkins and Van Kerm (2006).
[3] Consult Appendix B for raw Gini coefficients and bootstrapped standard errors.

1980 and 2017, regional income gaps grow by roughly 50 percent. Correspondingly, income growth is strongly pro-poor in the initial period, before turning regressive in the subsequent period. As compared with 1940-1980, there is considerably more rank turbulence or leapfrogging in the post-1980 period, implying that a new pattern of winner and loser regions was set into place in the 1980s.

In Panel 2 of Table 1, finely granular results confirm that each of the two major periods observed in Panel 1 are internally homogeneous, and externally differentiated. In each decade in the first period, income gaps are decreasing, growth is pro-poor, and there is only a modest amount of re-ranking. From 1980 forward, the reverse tendencies are evident, with 2000-2010 representing somewhat of an exception, one that might be explained by the Great Recession. The 1980-1990 period is of particular interest, in that it signals large proportional expansions in income gaps, leapfrogging and, though to a lesser extent, pro-rich growth in relation to other decades. The post-1990 decadal changes remain consistent with 1980-1990 in terms of signs, magnitudes after 1990 are comparatively modest, but the shock seems to have long-lasting cumulative effects.

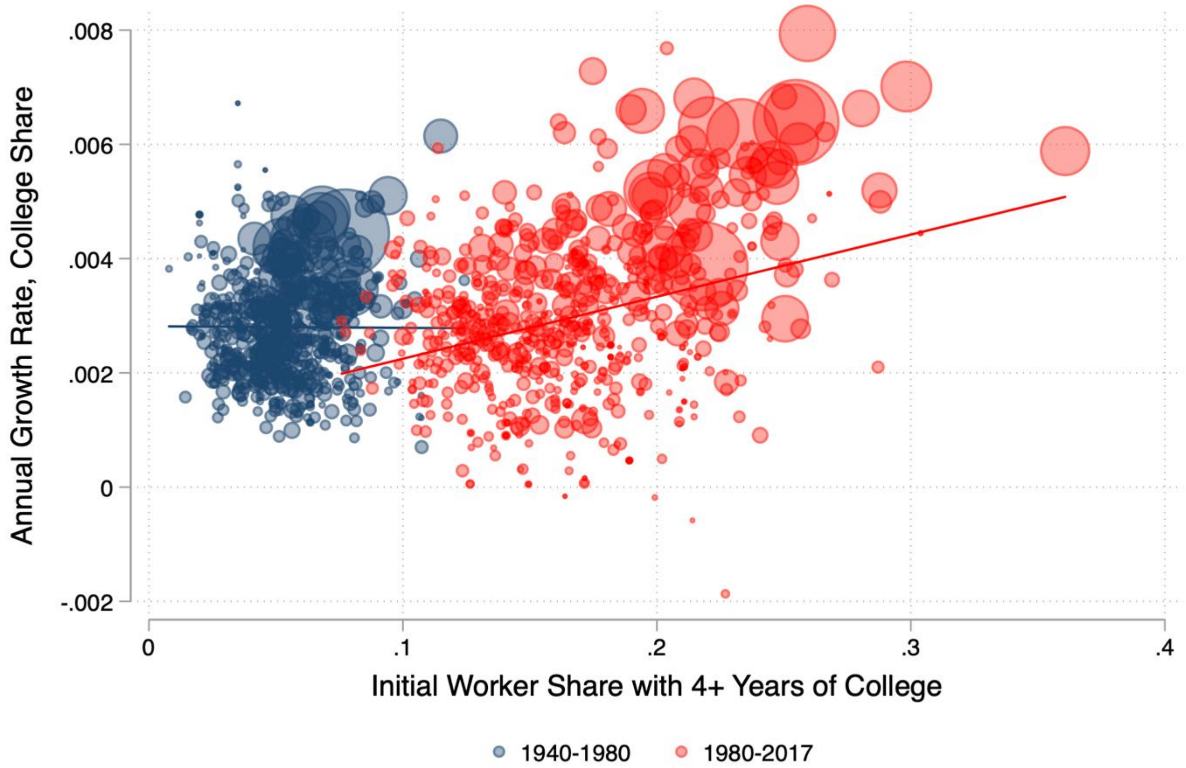
2.3 College graduates have concentrated in space since 1980, but were not doing so from 1940-1980

Studies of income inequality attribute a strong role to relative wage growth among the skilled, who are typically defined as the roughly 30% of the population that has attained four or more years of post-secondary education. Interregional inequality could grow without a general increase in interpersonal income inequality, but in the period since 1980, the two are intertwined, and divergence is therefore structured by differences in rewards to educational attainment and their changing geography. We consider each in turn.

Figure 2 relates initial local shares of college graduates with their subsequent annual growth rates. Starting with the more recent period, in red, the upward sloping relationship between initial shares and subsequent growth means that locations with better educated workforces in 1980 improved their endowments of college graduates more than places with weaker initial shares of college graduates. Internal and international migration, as well as in-situ factors have reinforced initial educational advantages. Prior work, such as Moretti (2004) and Diamond (2016) find similar patterns for metro areas over the 1980-2000 period. An urban size effect is also evident, in that it is largely initially more-populous cities whose shares of college graduates have grown the most, which means even larger growth in absolute terms, a result confirmed by Davis and Dingel (2018). Size and skill are closely linked in the reproduction of inter-regional economic divergence today.

The pattern for the 1940 to 1980 period, in blue, suggests these relationships were not in evidence during the period in which interregional incomes were converging. During this period, skilled workers were not concentrating in initially skill-abundant locations. The fairly flat linear fit line suggests no clear pattern, and while growth rates appear somewhat higher for larger cities, this growth is not evidently linked to initial skill endowments.

Figure 2. Initial shares of workers with 4+ years of college (blue=1940; red=1980) and subsequent annual growth rates in the share of workers with 4+ years of college, U.S. Commuting Zones.



Note: Circles represent 722 Commuting Zones, scaled according to population in 1940 (blue) or 1980 (red). Dashed lines represent each linear fit. 'College share' on the *y*-axis is the share of the working population who have completed at least four years of College education. Source data are public-use (IPUMS) extracts of the Decennial Census and American Community Survey. Full details of the data in the Data Appendix.

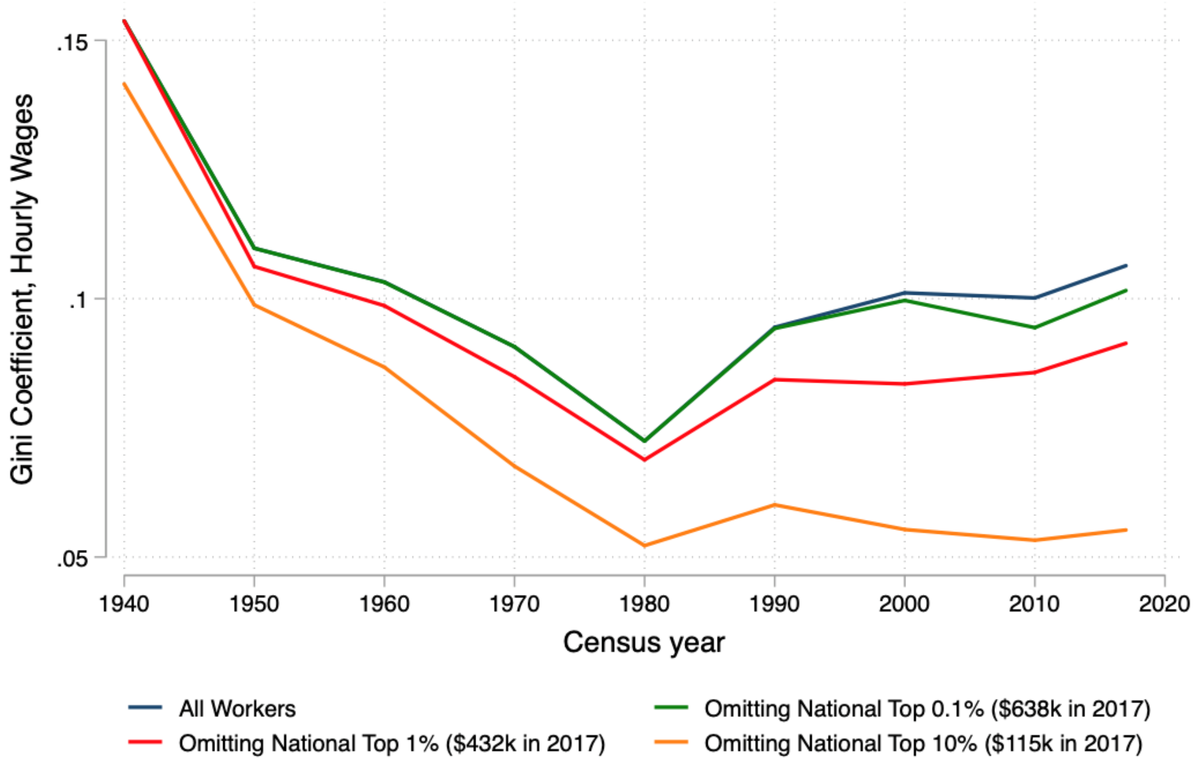
2.4 Divergence is not just about the one-percenters, but about the college-educated

Changes in interregional inequality have coincided with major growth in national interpersonal inequality, in particular the expansion in the income and wealth shares held by the top 1 and 0.1 percent of the population (Piketty and Saez, 2003; Alvaredo et al., 2013). Could the growth in interpersonal top income shares explain the recent rise in regional inequality? To consider this, Figure 3 again visualizes changes in regional Gini coefficients, this time omitting workers above the top 0.1, 1 and 10 percentiles of the national income distribution.

For comparison, the blue line in Figure 3 revisits the blue line in Figure 1: it measures interregional income inequality for all in-sample workers. The green line in Figure 3 indicates the extent of interregional inequality that remains when we remove individuals earning above the top 0.1 percent of the national income distribution. It largely mirrors the pattern for all workers, indicating that top earners are not the primary determinant of the recent rise in regional income inequality, though their importance has somewhat increased recently. The red line, which excludes those above the top one percent, indicates a moderately lower level of divergence after 1980, while the yellow line, which excludes workers above the top tenth percentile, remains roughly flat after 1980. Overall then, Figure 3 confirms the observation by Manduca (2019) that regional divergence since 1980 is explained largely by workers above the 90th percentile of the national income distribution. But it is not principally driven by Jeff

Bezos or the Koch brothers. Materially, cutoffs in terms of annual wages in nominal 2017 dollars for the top 10%, 1%, and 0.1% are, respectively: \$115,000, \$432,000, and \$638,000. The top 10% and even the top 1% include a lot of high-wage professionals and successful business owners. We still get a notable shift from interregional convergence to divergence when we omit workers far below the incomes of the very richest individuals in the United States. Equally, Figure 3 illustrates that much of the rise of divergence is driven by workers with incomes between the 90th and 99th percentiles of the national income distribution.[4]

Figure 3. Evolution of interregional income inequality (σ-convergence), overall versus specific percentiles of the national income distribution, 1940-2017



Note: Based on year-specific Gini coefficients estimated using average estimated hourly wage and salary income for all workers in 722 1990-vintage Commuting Zones (CZs), and weighted by population. Incomes are adjusted for inflation to 2015 dollars using Bureau of Labor Statistics CPI. Source data are public-use extracts of Decennial Censuses and the American Community Survey. Non-solid lines omit workers above a particular income threshold. For instance, the dotted line captures convergence patterns among workers who earn below the 90th percentile of the national income distribution. Further details in the text and in the Data Appendix.

[4] Income data in Census extracts are top-coded to maintain confidentiality. To the extent that high-income individuals are concentrated in space, then Figures 1 and 3 will be biased, but in the same direction. This means we may be understating the true extent of regional disparities, but it does not challenge the fact of divergence, or the role of certain income thresholds in it. Capturing wealth rather than income could also lead to higher estimates of disparities. We are unable to accurately measure the changing geography of wealth.

2.5 The returns to a college degree have become increasingly place-specific since 1980; they were becoming ever-less so between 1940 and 1980

Considering the location-specific returns to skill, we return to Figure 1. The red line traces the evolution of the Gini index for hourly pay across CZs for college graduates only, while the green line captures the pattern for those with less than 4 years of college. These lines tell quite different stories. Consistent with results on β-convergence among metropolitan areas from Giannone (2017), we find that the end of income convergence occurs only for workers who hold college degrees: the recent turn to divergence is driven by increasing inter-place inequality *among* college graduates. This suggests place-based dynamics at work: after 1980, the returns to higher education become stratified by location: the wage premium for college workers rises generally, but there is a strong gradient related to city size and density (Autor, 2019). Contrastingly, for workers with less than four years of college, except for a small interruption between 1980 and 1990, interregional income gaps have declined in each decade. Wages for these workers have converged over geographical space over the nearly 80-year study period.

2.6 Superstar cities remain super after accounting for housing costs, but the pains of higher prices are felt strongly by workers without college degrees

Recent income growth has been in more populated and skill-abundant regions; high local prices, and in particular housing costs could offset higher salaries for the college educated, and reduce real incomes divergence overall (Moretti, 2013). The high cost of housing is widely considered to be among the most important urban issues of our time (Hsieh and Moretti, 2019; Glaeser and Gyourko, 2018; Anenberg and Kung, 2018; Rodriguez-Pose and Storper, 2019).

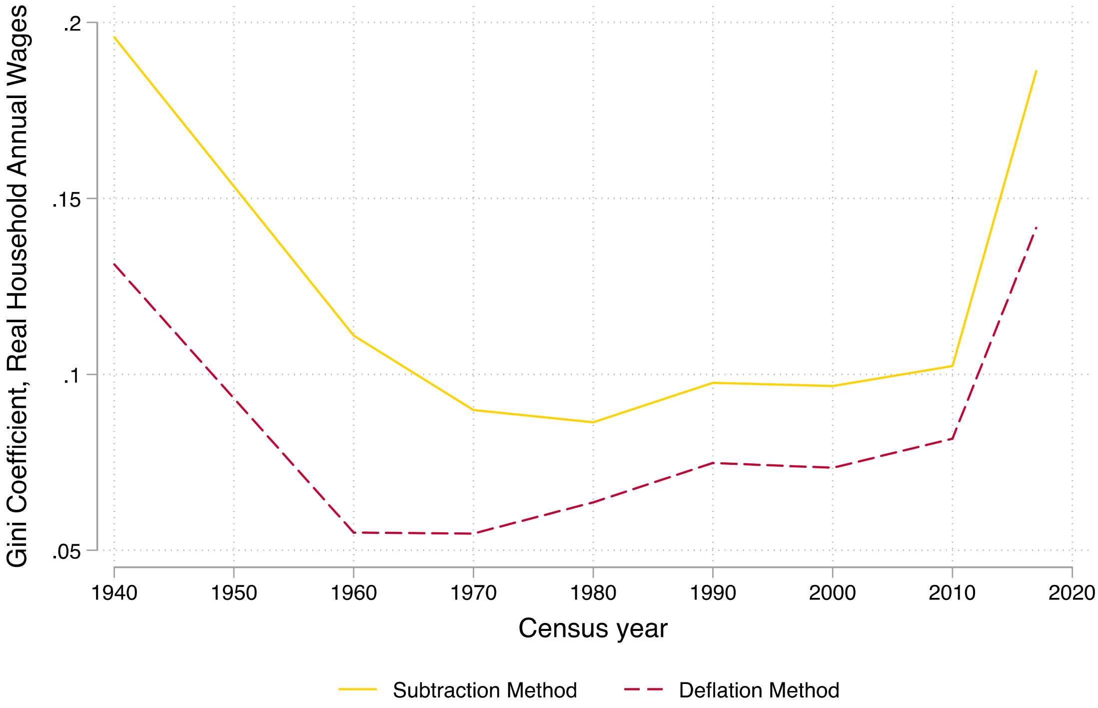
To enable this kind of comparison, we report the evolution of 'real' (housing cost-adjusted) household incomes in Figure 4. Lacking long-run representative information on local differences in non-housing costs, we follow common practice in focusing on the price of housing (i.e. Moretti, 2013; Kemeny and Osman, 2018). Housing consistently makes around up 40 percent of the Bureau of Labor Statistics' urban consumer price index (CPI-U), and it varies across regions far more than most other consumer costs. We estimate real income using two methods. First, to capture effective take-home pay net of housing, we directly deduct reported (or for owners, imputed) rental costs from household income. For renters, rents consist of annualized reported monthly gross rent. For owners, rents are imputed as the median annual rental costs for households that are analogous in terms of a combination of commuting zone of residence; number of rooms in the dwelling; and household maximum educational attainment.[5] We estimate median real wages for each location for the full sample of households. Our second method resembles that described in Moretti (2013), in which for each commuting zone and year median nominal wages are deflated using median rents.

For both accounting methods, real wage inequality evolves in a manner that resembles nominal wage inequality. Broadly, after accounting for living costs, interregional inequality declines and then rises. The inflection point does depend on the approach taken, with the shift to divergence occurring one decade earlier using the deflation method. Overall though, Figures 1 and 4 resemble one another. Quantifying this resemblance, the correlation

[5] For this purpose, educational attainment is categorized as one of four mutually exclusive categories: less than a high-school diploma; high-school graduate; some college; and at least four years of college.

coefficient between median local nominal and the different real income methods is not less than 0.72 in any year, indicating that today's nominal superstars largely remain super after accounting for living costs.

Figure 4. Evolution of 'real' interregional income inequality (σ-convergence), subtraction and deflation methods, 1940—2017



Note: N=722 Commuting Zones. Based on year-specific Gini coefficients estimated using average household real annual hourly wage and salary income for all in-sample households in 1990-vintage CZs, weighted by population. Source data are public use extracts of Decennial Censuses and the American Community Survey. Further details in the text and in the Data Appendix.

Along with a range of less populous locations, the top five percent of real incomes in 2017 features commuting zones centered on cities like Boston, San Jose, Houston, Los Angeles, San Francisco, Dallas, and Salt Lake City. The presence of these superstars near the top of the list does not mean that productivity and demand are not capitalized into housing markets, but that nominal income growth in some of these cities – notably places like Boston and the Bay Area – has outstripped the upward expansion in the cost of housing. This fits evidence from Davis and Ortalo-Magne (2011) which finds that places like San Francisco are in fact less expensive that they 'should' be.

In sum, the magnitude of post-1980 interregional wage inequality is only somewhat diminished when we account for housing costs. In sections 3 and 4 we will argue that housing cost gaps emerge in divergence periods due to the spatial distribution of the skilled, while in convergence periods they diminish because of the spread of the skilled. In divergence periods, due to non-homothetic preferences and selective migration patterns, they have little independent influence on real income divergence, contrary to the findings of Hsieh and Moretti (2017).

2.7 Divergence in total utility is multi-faceted and possibly greater than real income divergence

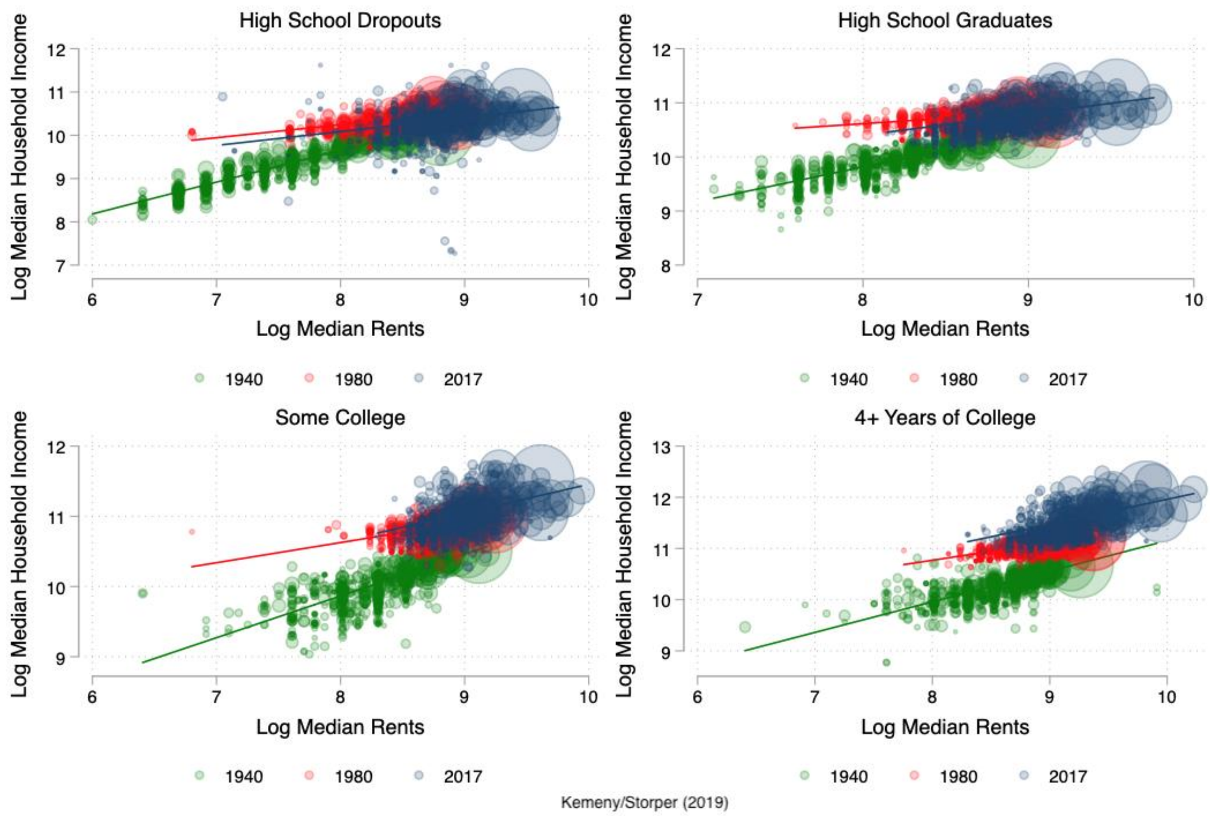
Part of the reason for the limited causal role for housing in divergence periods observed above is that in such periods unskilled have limited job opportunities in prosperous centers (no matter what their housing supply), and the skilled prioritize more and different access to amenities than the unskilled (non-homothetic preferences). Non-housing costs not only do not reduce the observed advantages of large superstars, they reflect the additional advantages of these places. For instance, using quality- and variety-adjusted accounting of an inclusive basket of consumer items, Handbury and Weinstein (2014) find that city-size and prices are negatively related, echoing the classical NEG model of polarization where consumer prices are lower and consumer variety higher with urban scale.

Nominal and real income divergence could fail to fully capture patterns of some wider notion of welfare inequality to the extent that wages and prices leave unmeasured the uneven availability of amenities like pleasant weather or cultural institutions. The nature and location of second-nature amenities are largely endogenous, driven by induced local demand from local incomes and the tastes and preferences associated with lifestyles of people in different types of occupations (Diamond, 2016). High-income locations are well endowed with amenities (Kemeny and Storper 2012). Couture et al (2019) go further to argue that, because we cannot index the utility value of endogenous amenities, the average real utility of high-skilled workers in high-skill cities may be underestimated by real incomes. A telling indicator of the appeal of amenities and employment opportunities in today's superstar cities is the time that each generation spends living in inner metropolitan areas, reducing the volume and time span of suburban outmigration in child-rearing years (Lee et al, 2019).

Moreover, larger and more skilled cities may offer greater opportunities for learning and experience accumulation (De la Roca and Puga 2017; Glaeser and Mare, 2001). Higher rates of job turnover and interpersonal interaction in cities spur experiential learning effects that, through the matching process, are then capitalized into steeper wage growth across the work life cycle. The benefits plausibly increase the benefits for the skilled of locating in certain cities, and further differentiate them from locations that do not offer equivalent matching dynamics. Viewed from a life-cycle perspective, a recent college graduate may face painful housing choices in superstar cities at the beginning of the career, but the long-run income trajectory will be superior than if they pursue a career in a smaller and cheaper and less skilled locale. This works in reverse as well: individuals in other locations may be progressively locked out of networks that would enable them to match to jobs in certain locations, assuming they could acquire entry-level (college graduate) skills in the first place.

While the preceding evidence suggests that the median household in many superstar cities is well off in real terms, such locations may not offer this surplus to everyone. Ganong and Shoag (2017) observe that the average janitor in New York City has experienced falling real wages relative to an occupational counterpart in the parts of Southern U.S. Figure 5 assesses the generalizability of this for households in each of four education categories. Defining each household's education on the basis of its most-educated member, Figure 4 relates changes in the relationship between median annual household nominal wages and annualized median rents for each education class. For readability, we present results for only 1940, 1980 and 2017. An upward-sloping relationship indicates that workers living in higher rent locations are compensated for these higher costs with larger incomes. Each linear fit line is upward sloping, yet the top two panels of the figure indicate that for workers with a high school diploma or less, compensation for living in costly locations has considerably declined over time. For workers with some college, the slope of the relationship has declined moderately. Finally, for workers with four or more years of college, the relationship remains largely unchanged since 1940. There is a real interpersonal inequality dynamic layered within the broader patterns of interregional inequality.

Figure 5. Median household nominal incomes and median rents, 1940-2017



Note: N=722 Commuting Zones. Median rents and incomes are specific to each year, education category and Commuting Zone. Linear relationship is weighted by population. Rents and incomes are logged. Source data are public use extracts of Decennial Censuses and the American Community Survey. Further details in the text and in the Data Appendix.

2.8 Superstar cities are highly internally unequal.

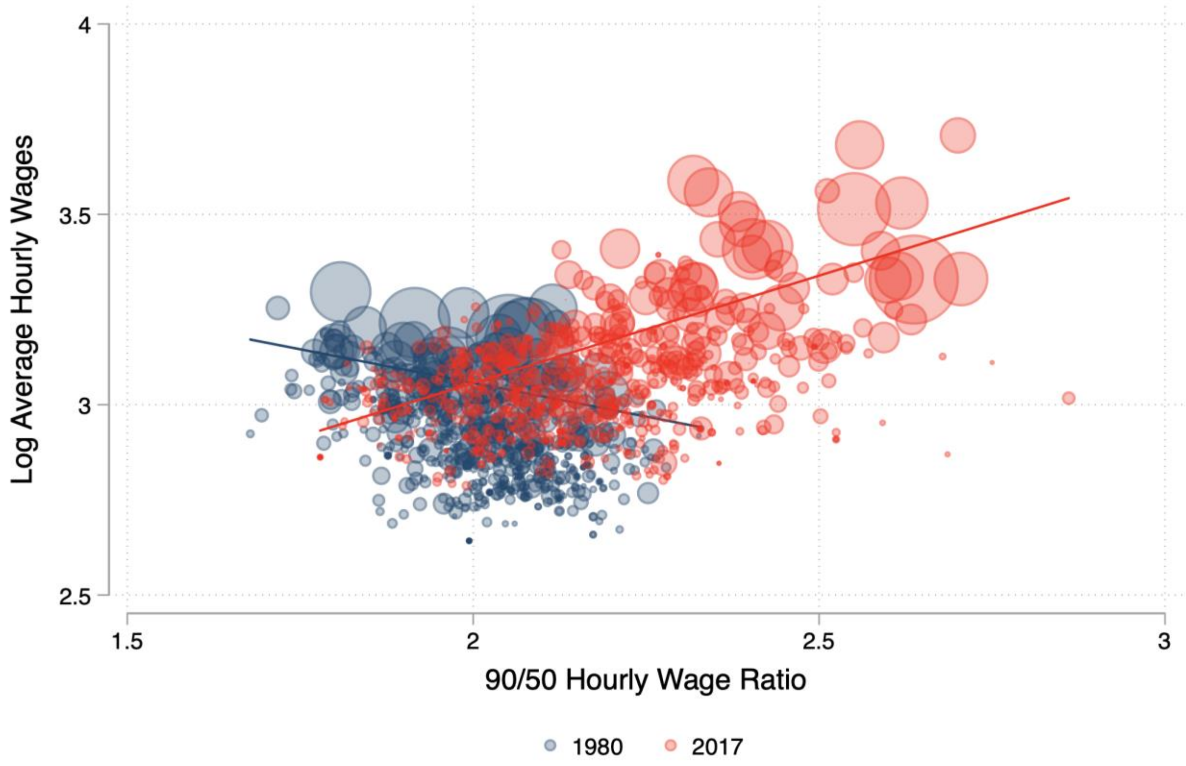
Researchers increasingly recognize that today's high-income cities also feature high levels of interpersonal inequality (Abel and Deitz, 2019; Florida, 2017). Figure 6 confirms this by plotting the relationship between local nominal average hourly incomes and inequality in 2017 and 1980. Inequality in this figure is measured as the ratio of incomes at the local 90th percentile to those at the 50th percentile.

A few points stand out in Figure 6. First, the largest wage ratios are absolutely larger in 2017 than 1980, which means that the most unequal cities today are more unequal than the most unequal cities in 1980. Second, in 2017, the most unequal cities also tend to be larger, a pattern that is not present in the 1980 data.[6] Third, and most strikingly, the fundamental

[6] These three patterns are robust to the use of annual rather than hourly wages. They remain materially similar when 90/10 wage ratios are used to measure inequality. Further, results do not depend on the log transformation applied to the wage data. Pre-1980 series resemble 1980, and are available upon request.

nature of the relationship between incomes and inequality has reversed. While the contemporary series confirms that richer places are the most unequal today, the 1980 series, in blue, is downward sloping, indicating that locations that are more unequal are also those with lower average wages. This pattern is true not just for 1980, but for each decade of the convergence period.

Figure 6. Log hourly wages and 90/50 wage ratios, U.S. Commuting Zones, 1980 and 2017



Note: N=722 Commuting Zones. Incomes are adjusted for inflation to 2015 dollars using Bureau of Labor Statistics CPI. Circles are scaled according to population in a given year. Linear relationship also weighted by population. Source data are public-use extracts of Decennial Censuses and the American Community Survey.

This shift in patterns of inequality also relates to evidence on the links between nominal wages and rents (Figures 3 and 4). Putting this together, we conclude that the greater real income inequality in high-income, graduate-abundant cities is driven by a combination of the spillover effects of high incomes on housing markets and the limited elasticity of their housing supply (Hsieh and Moretti, 2017; Rodriguez-Pose and Storper, 2019).

2.8 Over the very long-run, the recent turn to divergence is not an isolated occurrence

If skill-biased technologies of the third industrial revolution might be driving the current divergence, it makes sense to consider the effects of prior industrial revolutions as well. This could help determine whether the contemporary switch is an idiosyncratic one-time event, or if there is instead a deeper, technological logic at work.

While data on the geography of the period around the first industrial revolution is unavailable, second industrial revolution consisted of the widespread application of mechanical engineering technologies, driven by fossil fuels and electricity, to a wide variety of domains. Commonly believed to have begun in the 1860s, the revolution placed capital-intensive, large-scale, electrified manufacturing at the center of the economy. Like today, it had a significant skill bias, compared both to previously existing manufacturing and to work in general. A sensible equivalent in the second industrial revolution for today's college graduates would be manufacturing workers. A good proportion of these workers, particularly machine operators, would have been skilled relative to the economy as a whole, and especially in contrast to agricultural workers, who composed a still-important part of the workforce in the 1880s.

Table 2. National Estimates of the Ratio of Manufacturing to Agricultural Wages, 1890-1945

Year	Manufacturing Earnings/ Farm Earnings (Annual)	Manufacturing Earnings/ Farm Earnings (Monthly)
1890	1.88	-
1895	1.93	-
1900	1.76	-
1905	1.64	-
1910	1.66	1.85
1915	1.60	1.92
1920	1.68	1.71
1925	2.18	2.51
1930	-	2.47
1935	-	3.59
1940	-	3.59
1945	-	2.16

Note: National wage data transcribed from annual summary of statistics of the U.S. Census Bureau. The authors thank Sergio Petralia for sharing these data.

Capturing the geography of incomes for this period remains a challenge. The Census Bureau first recorded Decennial respondents' wage information in 1940. Prior to this there is no direct way to capture the evolution of average wage levels at the scale of commuting zones, metro areas, or counties. We try to work around this limitation inferentially. As we have seen, today's divergence is largely due to the wages, wage premia, and spatial concentration of the college graduates, who are the workers performing tasks that are most complementary to the technologies driving the third industrial revolution. As a first point, then, we argue that understanding wage gaps in the second industrial revolution requires measures of the location and relative rewards enjoyed by manufacturing workers. Table 2 provides some information on relative rewards. It shows that, between 1890 and 1945, national average manufacturing wages were between 1.6 and 3.6 times higher than those of agricultural workers. We make a direct analogy between this and today's skill premia.

Regarding the geography of these workers, we exploit information from Historical, Demographic, Economic, and Social Data: The United States (Haines, 2005), which compiles information from Decennial Censuses, the Census of Manufacture, and other sources, providing information for States, and crucially for our purposes, for counties. We gather county-level data on manufacturing output, capital investment, average wages, and employment counts from 1860 onward, with the longest series tracking workers.

Figure 7 shows the Gini coefficients for the geography of these series. For comparison, in blue we also include coefficients that track annual wages at the Commuting Zone level for 1940 to 2017. The earlier series provides some sense of the spatial distribution of the high-wage work (manufacturing) that was strongly complementary to the Second Industrial Revolution. Measured in terms of employment, capital, and output, manufacturing activity concentrates after 1860 and then remains at a high level – with several peaks on the mountain range – until about 1940. At that point then, for the employment series for which more recent data is available, we can observe a long-term spreading-out that follows a similar trajectory to the income convergence for all workers tracked in the ‘modern’ series up to 1980. We can think about this in relation to the more recent empirical evidence in this paper. As manufacturing work becomes spatially dispersed in the post-war period, we observe strong interregional convergence of wages for both skilled and unskilled, and no strong relation between presence of the skilled and in-migration of the skilled. Although inferential, this looks very much like the spatial diffusion of the second industrial revolution, after an initial concentration of that revolution.

Figure 7. Long-term geographical patterns of convergence and divergence, 1860-2017

Note: Historical series from 1860 drawn from data from ICPSR 2896 (Haines, 2005). Unit of observation in these data is the county. Each Gini in these series uses the left Y-axis. On the right Y-axis is average annual wages in constant 2015 dollars measured among Commuting Zones using Decennial and ACS data.

This contrasts to the standard view in the convergence literature, which is that the post-war period is part of a long-term process of American convergence, from colonial origins to economic maturity and greater geographical integration. From the mid-19th to mid-20th century, the underlying secular forces for convergence of the American space economy were very strong (Pred, 1973). Following the Civil War, the U.S. continued to occupy the West and admit more states, to improve long-distance transport by completing the trans-continental railroad system, as well as major navigation improvements on major waterways, and – especially in the post-1890 period, improvements in communications. All of this was coupled to high overall rates of inter-state labor mobility (Ganong and Shoag, 2017).

Given these long-term forces for integration, the two episodes of spatial income polarization are all the more striking and suggest that other forces were at work. This point of view is buttressed by the occurrence of another wave of interregional income polarization today, in a country that has completed its frontier migration transition and has nearly complete infrastructure coverage and low transport costs. Stated another way, the spread of infrastructure like the Interstate Highway System appeared to drive convergence in the 1950-1980 period, but that same infrastructure is apparently compatible with convergence reversal after 1980. Perhaps the highway system was actually the weak force for convergence from 1940 to 1980, but its effects were masked by our failure to consider the spatial diffusion of the maturing second industrial revolution, just as today it is insufficient to generate convergence faced with a spatially concentrated and still incomplete third industrial revolution. In other words, convergence models should not infer causes by examining the limited sample of evidence that comes from phases of convergence, if in fact convergence is only temporary.

3. Explaining the alternation from convergence to divergence: Supply- and demand-side theories

Most regional economic models are marked by expectations of convergence (Borts, 1960). Classical versions are abstract, concentrating on production functions involving homogeneous capital and labor. Spatial equilibrium theory made a great step forward, allowing for different types of capital, labor and preferences, with an emphasis on factor mobility (Rosen, 1979; Roback, 1982; Glaeser, 2008). In this view, household locational arbitraging generates the tendency toward convergence of real wages or, in a more expansive version, total utility. Rather than nominal incomes, for comparable workers at the margin at least, real incomes or utility ought to equalize. These are thus supply-side models with individual or household preferences for wages, rents and amenities driving convergence processes.

The empirical reference point for this body of work is the mass migration of labor and capital in the U.S. between the 1950s and the 1970s. Workers' locational choices are said to be the foundational driver of these phenomena, driven by their preferences for natural amenities like mild winters (Rappaport, 2007), as well as for cheaper, bigger, lower-density housing (Glaeser and Gottlieb, 2009). In turn, these choices were allegedly enabled by shocks to infrastructure after the 1940s, notably better highways, and the wide availability of air conditioning (Glaeser and Tobio, 2007). Concretely, an increasing proportion of workers were said to choose cheap suburban living, lower wages and more sunshine, while a declining proportion maximized nominal wages but paid higher rents and enjoyed less pleasant winters. On balance, however, both skilled and unskilled workers were spreading out during that period, so that regional gaps in the composition of the skilled and unskilled were declining, and regional returns to education were converging.

A different set of preferences and constraints appear to motivate the distinctive sorting patterns of the post-1980 period. Two main supply-side explanations are adduced to explain the behavior of college graduates. One is that they prefer co-presence due to advantages in production (Glaeser and Resseger, 2010). College graduates make each other more productive when they are co-located (Moretti, 2004; Shapiro, 2006; Davis and Dingel, 2018). Rates of skill and experience accumulation are higher in more urbanized and skill-abundant places (Glaeser and Mare, 2001; De la Roca and Puga, 2017). These patterns offer support for the wage gains that college graduates have enjoyed in skilled cities. A second possibility is that graduates are making locational choices to maximize the utility they derive from access to amenities. Specifically, they may prefer humanly generated amenities that are found in centers of large, dense cities (Florida, 2002; Chen and Rosenthal, 2008; Moos et al., 2018; Lee et al., 2019; Couture et al., 2019). These amenities largely depend on local incomes (Diamond, 2016; Couture et al., 2019), setting off a snowball dynamic of further attractiveness to the skilled. Behind these two non-exclusive possibilities, a common assumption is that today's college graduates have locational preferences for the urban that do not resemble those of their suburban parents.

To make the link to divergence, supply-side arguments emphasize barriers to migration. In these accounts, high and fast-rising housing costs in high-income cities, driven by inelastic land supply and increasingly restrictive land use regulation, have inhibited less-skilled workers' mobility (Gyourko et al., 2013; Ganong and Shoag, 2017). Relative to expected wages, workers without university degrees find housing costs in superstar cities to be prohibitive. This deters their in-migration. Extending this logic, Hsieh and Moretti (2019) argue cheaper housing in superstar cities would powerfully reshape the U.S. urban system, transforming the current distribution of population and productivity. The principal assumption behind this prediction is that left-behind regions contain deep pools of workers who would actually prefer to be in skilled cities; in their simulations of an urban system with lower housing costs, major increases in migration from less prosperous to more prosperous regions take place. Today's largest metropolitan areas would become much bigger and considerably more productive.

As we shall see, these accounts are not fully convincing. But even if they were, they would leave unexplained the mechanisms that flipped the system from convergence to divergence. To do so from a supply-led perspective, we require an explanation of the change around 1980 in workers' preferences for wages, interactions, or amenities, or for that matter, of the sources of the different preferences held in the previous period.[7] Thus, we would be able to explain why skilled workers in the late 19th century preferred co-location in dense manufacturing cities, but then in the post-war period they came to prefer sprawl, sun, and cheap housing, and then again after 1980 the preferences held by skilled worker switched back again towards spatial concentration. Even if such switches are in the realm of the sociological, any supply-led model of regional development rooted in individual preferences has no causal purchase unless it can deal with the origins of preference change. Equally, models that emphasize migration barriers for the less skilled must account for why land use laws had such different

[7] One attempt to reconcile these apparently different preferences has to do with falling urban crime rates (i.e. Schwarz et al., 2003; Glaeser and Gottlieb, 2009). In this vein, skilled workers might (always) prefer cities, but high crime rates deterred them from acting upon these preferences. One challenge with this argument is that urban crime began to decline during the 1990s, while we know that the shift in income inequality begins a decade earlier.

effects prior to 1980, or at least convincingly argue how new restrictions on supply interact with increasing demand for city locations. Considering the mobility of less-skilled workers, while Ganong and Shoag (2017) do cite a literature positing a role for race, politics and other factors, these are not clearly exogenous to incomes. Nor does the timing of changes in land use regulation link neatly to changes in the dispersion of incomes, with some evidence suggesting a new regime begins in the early 1960s (Garrett, 1987; Fischel, 2004).

Our alternative view starts with the geography of labor demand. In its simplest form, the third industrial revolution has reshaped the nature and geography of firms' demand for workers. Computers and related new technologies made college graduates more productive and, in combination with offshoring, they replaced a great deal of routine middle-skill work, (especially since the post-2000 so-called China Shock). The demand for skilled workers also became increasingly geographically concentrated. Hence, workers are not making choices to satisfy their desires for sunny Januaries, yoga studios or co-presence; the fact is that the map of destinations in which skilled workers can find suitable work has been reshuffled and it has overwhelmingly concentrated in large cities.

A wealth of studies supports these ideas. Lin (2011) documents how new occupations that emerge as a response to technological change are concentrated in skill-dense urban environments. In a similar vein, Berger and Frey (2016) find that the computer revolution after 1980 spurred the creation of new nonroutine-intensive occupations, which were strongly spatially agglomerated. Other researchers seek to capture related insights in a general equilibrium context. Diamond (2016), for instance, develops a model whereby skill-biased technological change shifts firms' relative demand for higher- and lower-skilled workers. Initial endowments of skilled workers mean that cities are rewarded differently by the shock, with such advantages then becoming self-reinforcing, partly through the endogenous provision of amenities.

But others argue that on average the successful regions that pulled away from the rest did not do so merely because of their pre-1980 skilled labor endowments. Thus, Baum-Snow and Pavan (2012, 2013) and Baum-Snow et al (2018) argue that the technology shock doesn't build on pre-existing factor complementarities, it reshapes them through an increasing factor bias of agglomeration economies, growing industry-specific skilled labor pools and specific employer pools together. This lines up well with Autor's (2019) finding of rapid rises in the returns to skill in cities, and both find a considerable and increasing size premium. Baum-Snow and Pavan (2018) estimate that big cities contribute at least 25% of the total increase in national wage inequality, and of this, about 80% is due to the rising factor bias of agglomeration economies. This view is also consistent with the notion of an endogenous local and regional component to technological change or adoption (as in Acemoglu, 2002; Beaudry, Doms, Lewis, 2010) and with experience effects in agglomerations (De la Roca and Puga, 2017).

Going further, Milanovic (2019) argues that top incomes today involve a greater overlap between labor income and capital income than in previous periods (such as the early 20th century), as some of the skilled accumulate substantial capital from their high labor income. This, in turn, is reinforced through the rise in assortative mating, in which educated 'power couples' have an increasing urban bias because cities offer both partners a greater probability of successful skills matching (Costa and Kahn, 2000). All of these are recent dynamics that could be shaping the attractiveness of large cities to the skilled in a way that was not the case prior to 1980.

In spite of all this progress in both identification and in contextual history, with the notable exception of Giannone (2017), there are few attempts to capture both convergence and divergence in a single framework. To do so, requires an explanation not just of recent divergence but of alternations in labor demand and its geography.

Consider how such intuition prompts a reconsideration of the Rosen-Roback-Glaeser supply-side narrative of the pre-1980 period. Instead of workers looking for sun and sprawl, such a framework starts with the second industrial revolution that began in the 1880s. That revolution, centered on electrical-mechanical technologies, at first concentrated its supply chains in Northeastern and Midwestern cities (Figure 6). Skilled (and unskilled) manufacturing workers concentrated there, drawn by the availability of relatively high-wage jobs. As the ideas powering that industrial revolution matured, skills involved in key production tasks became increasingly codified and embodied in machines. As the scale of industrial activity grew, unit transport costs within supply chains declined. This decline was further reinforced by the extension of transportation infrastructure such as the Interstate Highway System. The result was technology diffusion following an initial period of high geographical concentration of the revolution. This fostered the dispersion of manufacturing activity at the domestic scale, and eventually a wider global unbundling of supply chains (Baldwin, 2006). Firms found cheap land and a non-unionized workforce in the South, at first in the less-skilled parts of supply chains, and with time, in skilled or formerly-skilled activities. A good deal of evidence points to the primacy of jobs or demand over supply movements in that period (Blanchard and Katz, 1992; Greenwood and Hunt, 1989; Partridge and Rickman, 2003; Kemeny and Storper, 2012; Norton and Rees, 1979). Anecdotally, as early as the 1940s, the U.S. Congress expressed concern with the emergent shifts in the geography of labor demand: 1949 witnessed the publication of its commissioned report on “Why industry moves South” (McLaughlin and Robock, 1949). By the 1950s and 1960s, the industrialization of the South was well underway, with relative deindustrialization of many of the industrial cities in the Northeast and Midwest compared to the 1880-1930 period. Especially after 1945, the convergence that diffusion generated was further strengthened by long-run integration processes of the South.

Then, around 1980, a new industrial revolution shocked the geography of labor demand. This industrial revolution is centered on the microchip and emerging technologies enabling digital communication. These technologies give birth to entirely new sectors and completely transform many existing ones. The key innovation and management functions of their supply chains have elaborate divisions of labor and high spatial transactions costs, favoring agglomeration. The labor requirements of these activities are skill-biased, and in a context of rapidly changing skill and staffing requirements, the access to large, diverse skilled labor pools labor pools enhances labor matching and strengthens the productivity advantages of agglomeration. Demand-side work in spatial economics considers whether such agglomeration and urban wage premiums are due to sorting of firms or due to the matching of firms and other resources within agglomerations (Combes et al, 2008; Fontagné and Santoni, 2018). Whatever the precise combination of sorting and matching, the shift from convergence to divergence strengthened both of them.

The historical literature on technological change offers the missing causal element in the story. It suggests that, behind any significant changes in sorting or matching behavior of firms, is the degree of novelty of fundamental technologies and the frontier. ‘Technology’ here refers not to the gradual accretion of knowledge, but instead, using Mokyr’s (1990) evolutionary metaphor, to the kinds of mutations that generate entirely new species, that punctuate equilibria, and set the economy on a new path. These are what economic historians call ‘general-purpose technologies’ (GPTs), such as the steam engine, the electric dynamo, fossil fuels and associated mechanics, and the microchip (Bresnahan and Trajtenberg, 1995; Lipsey et al, 1998). Each of these GPTs ushered in an industrial revolution: the first around the 1820s, organized around textiles; the second starting in the 1860s and structured around mechanical and electrical engineering and fossil fuels; the third proceeding from the 1970s, dominated by information technology, finance and bioscience. Industrial revolutions act as a shock to development, focusing growth, entrepreneurship and wealth accumulation in the leading edges of the economy.[8]

There is also a line of thought that picks up on the geography of technology as a cyclical process, contrasting innovation with spatial patterns of concentration to maturation with diffusion (Vernon, 1966; Norton and Rees, 1979; Storper and Walker, 1989; Malecki, 2010; Myrdal, 1957; Pred and Hagerstrand, 1967;Roses and Wolf, 2018). Newer technologies raise the uncertainty and variability of markets, and this then raises the intermediate transaction costs of the sector, through the spatial costs of supply chains (sharing), increases in labor turnover and the costs of matching; and further technological innovation (Duranton and Puga, 2004). The factor bias of agglomerations tends to rise under these circumstances (toward cutting edge capital and skilled labor), and rent-sharing from high-productivity work and monopoly rents to innovative firms raise wages in industries that are strongly complementary to the new technologies. The result is the shift in the urban system from income convergence to divergence.

Revolutions ultimately become tomorrow’s routines, with codified knowledge more apt to diffuse away from initial innovating firms and industries, in turn bringing about a decline in technological rents to these firms and the categories of workers initially involved in the revolutionary activities (Schumpeter, 1939; Mokyr, 1991; Rosenberg, 1982; Crafts and Venables, 2003; Bourguignon, 2015; Perez, 2010). The maturation of each revolution causes general declines in skill premiums, but also through de-agglomeration, the spatial de-concentration of the skilled, reversing divergence and transforming it into a trend toward interregional income convergence.[9]

Once we incorporate the fundamentals of technological change in temporal-spatial perspective, a lot of the pieces of today’s puzzle of income divergence fall into place, but moreover, so does the previous period of convergence, which corresponded to a general de-agglomeration tendency as the technologies and industries of the second industrial revolution matured. It also aligns with a previous round of divergence in the late 19th and early 20th century (see Figure 6) – in the early part of the second industrial revolution – that to our knowledge remains unexplored in the economic geography literature. In any case, the

[8] There is a literature on technology and the fates of cities. But it tends to emphasize discrete technologies (such as computers), rather than industrial revolutions; and it assigns causality to adoption of technologies, rather than agglomeration of major innovation waves (Berger and Frey, 2016).

[9] We are not alone in incorporating technology shocks in spatial equilibrium frameworks. Glaeser’s (2008) synthesis of the Rosen-Roback-Graves perspective locates the causes of convergence in the technology shocks of the Interstate Highway System, air conditioning, and malaria control. As we point out above, significant movement of manufacturing to the South began well before the Interstate Highway System’s authorization, not to mention construction. “Smithian” revolutions in transport and trade technologies would predict a long-term steady movement toward convergence, however. They seem to be punctuated by a different sort of “Schumpeterian” technological change which, if powerful enough, can reverse long-term convergence. A full framework therefore needs to incorporate the two main types of technology shocks, as well as how they shape the production (labor demand) side and the labor supply (residential) side. Even in a fully integrated world, as long as there are some transport or communication costs that are positive, there are likely to be periodic episodes of divergence, especially as unit transport or trade costs rise in the face of uncertainty (innovation) or increases in variety (Duranton and Storper, 2008).

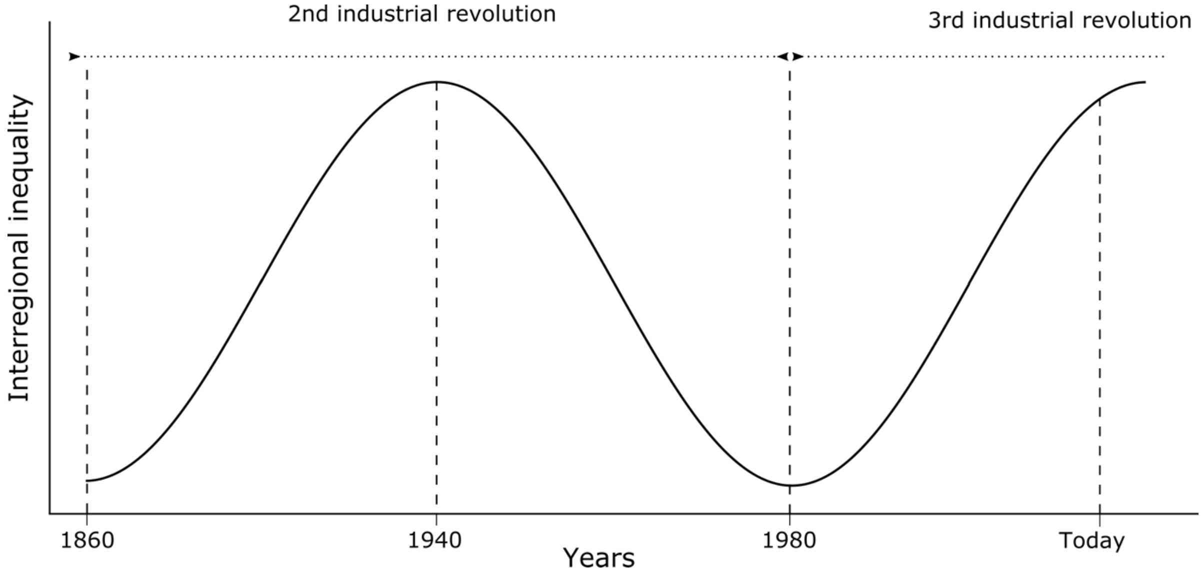
identification of the sources of today's agglomeration of firms and sorting of skilled labor and returns to such skills across space makes much more sense when placed in a framework that can explain the return of spatial concentration of skilled work after a long period of decline in the 1940-1980 period.

4. A demand-led model with periodic reversals

Synthesizing the insights of historians of technology, economic geographers and spatial economists, we consider that convergence and divergence represent distinct phases of a wavelike, historical dynamic of the geography of the economy. Alternations between divergence and convergence are regulated by technological disruption and diffusion, respectively. The arrival of a new industrial revolution punctuates existing technological and spatial equilibria. As new GPT technologies are initially localized and geographically selective, the new revolution concentrates demand and rewards, generating a pattern of divergence. The system eventually shifts to convergence as technologies mature and become routinized, allowing them to become organizationally and geographically diffused throughout the economy. This routinization and geographical dispersion tends to equalize the returns to both firms and workers across regions. This pattern continues until interrupted by the next round of disruptive innovation. In a stylized way, this wave pattern is depicted in Figure 8.[10]

We now outline the workings of a framework that actively incorporates both disruption and partial convergence or mean reversion. We do this in the form of a descriptive rather than formal model, yet our aim is to strip these two processes down in order to identify their core causal mechanisms.

Figure 8. Episodes of Interregional Convergence and Divergence



Note: Authors' elaboration

[10] There have been attempts to create an empirical predictive science of the timing of various kinds of waves in the economy, such as Kondratieff waves of output and investment. Though this work is clearly linked to general ideas developed by Kondratieff, we distance ourselves from predictive dimensions of that literature here, as we do not believe that the timing and precise magnitude of technology-driven waves of economic development can be anticipated, though they can be measured retrospectively.

4.1 Building Blocks

We start from an economy in which goods are produced using technology and labor. Technology, by which we mean general-purpose technologies, exist in two possible states: they start out being 'new', they later become 'old'. This aging process signals a progressive tendency from tacit to codified constituent knowledge. New GPTs are rooted in tacit knowledge, while the ideas embodied in old GPTs are fully codified. These potential spread or diffusion of a technology rises with codification.

Firms and workers also come in two forms: 'GPT-complementary,' or 'traditional'. GPT-complementary firms are those engaged in production in which the new general-purpose technology features importantly, while workers in this category have human capital that strongly complements the new GPT, whereby what constitutes skill lies beyond simply formal education and includes relevant tacit knowledge. Traditional firms and workers are those whose activities are chiefly reliant on the old GPT; their routines may also not be at all technologically-intensive.

The urban system is comprised of two locations, A and B, which are initially endowed with identical supplies of GPT-complementary and traditional workers. GPT-complementary workers are scarce relative to traditional workers in both locations. Returns to a given labor market segment are initially equalized across locations.

4.2 Assumptions

GPT-complementary firms and workers generate and benefit from scale-based localized productivity-enhancing externalities. In the context of technological rupture, the most important of these may be knowledge spillovers, but there can also be gains from what Duranton and Puga (2004) describe as 'sharing' and 'matching'.

We also assume that housing supply is less than perfectly elastic in A and B. This could be due to land use regulations that limit development, or because it takes time for developers to respond to demand by constructing dwellings. Further, A and B have an identical endowment of natural amenities and start with an identical supply of second-nature amenities (culture, sports teams etc). The supply of second-nature amenities is assumed to rise with local incomes.

4.3 The system in motion

By chance, the seminal innovations of a new GPT emerge in Region A. Some degree of GPT-complementary firm entrepreneurship in A is associated with this. [11] GPT-complementary firms in Region A reap the benefits of sharing, matching and learning, raising their aggregate output and productivity. New or expanding firms in this field consequently face incentives to locate in Region A, so that their advantages become self-reinforcing.

GPT-complementary workers in the urban system face the same centripetal forces. Those in B are incentivized to relocate to A where their specific skills are in high demand, and where they can benefit from matching and learning. GPT-complementary workers in A therefore

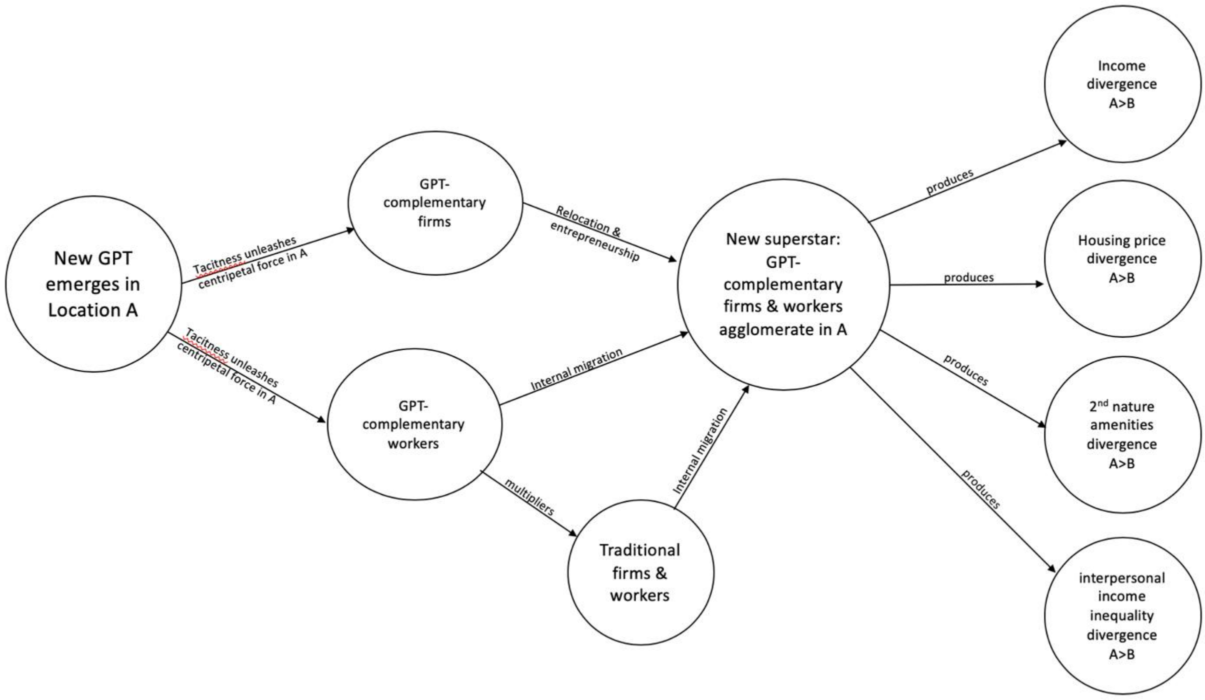
[11] In the real world beyond this very simple framework, 'chance' is not completely random. Innovation is a function of effort and endowments, the latter of which is highly unevenly distributed. Still, not all well-endowed locations perform equally well, as is clear from the comparison of Route 128 and Silicon Valley in Saxenian (1992), or San Francisco and Los Angeles in Storper et al., (2015).

earn more than their counterparts in B, through a combination of supply inelasticity, high human capital investments in skills, on-the-job learning, and technological rent sharing. Assuming some GPT-complementary workers remain in B, a skill-specific income divergence will be observed. Overall average wage levels in A and B also diverge. These patterns fit the divergence observed after 1980, and are consistent with aggregate patterns observed for the mid to late 19th century.

Local living costs in Region A rise with demand. However, monopoly rents – due to the rarity of their innovations, themselves due to tacitness, complexity and other agglomerative and economic features that limit their economic and spatial diffusion – enable a surplus that ensures that it remains profitable for revolutionary firms and skilled workers to remain inside the agglomeration. This is consistent with the earlier empirical observation that locations with high nominal incomes remain high in real terms, a pattern that is clearest for college graduates. Meanwhile, the creation of second nature amenities to serve highly-remunerated GPT-complementary workers further raises aggregate well-being in Region A relative to B.

Multipliers arising from the expansion of the revolution-intensive sector in Region A create jobs in non-tradeable sectors that employ traditional workers. Such workers may be local and also drawn from Region B. To account for higher local housing costs, nominal wages for these jobs are likely to be higher in Region A, however relative abundance means that the nominal wage premium will be lower than for GPT-complementary workers. Importantly, this premium may not persist after accounting for living costs.

Figure 9. The process of divergence



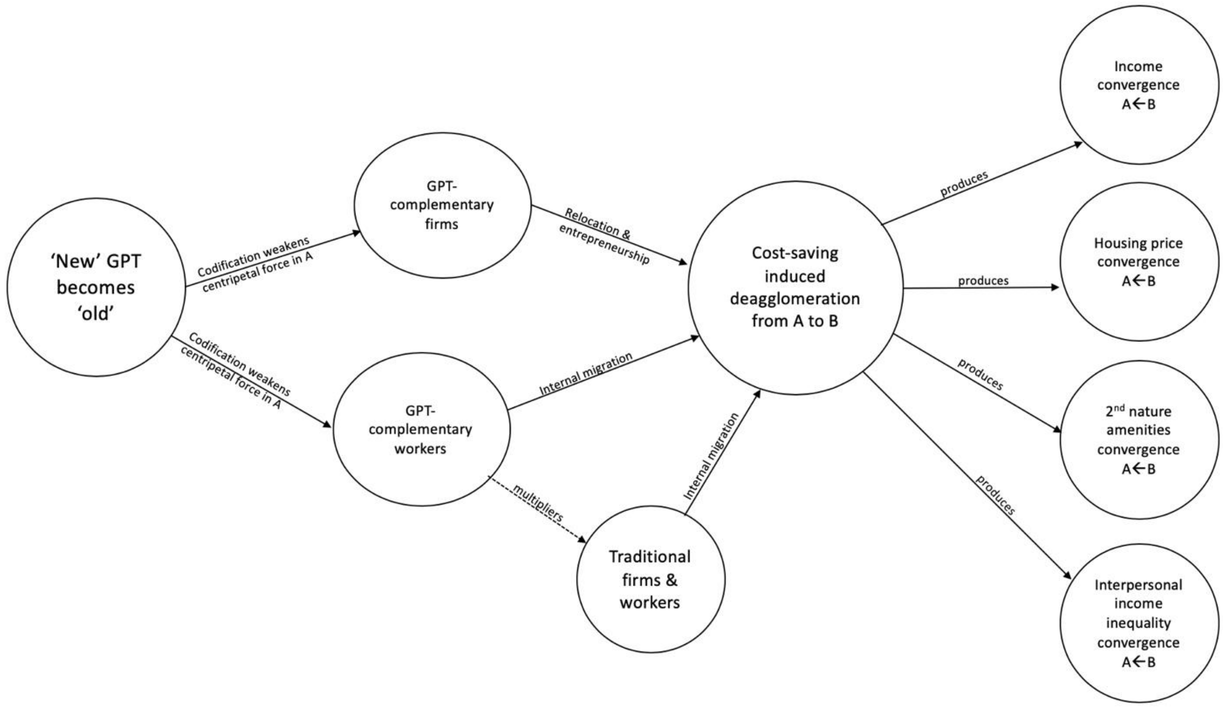
Note: Authors' elaboration

Consistent with the evidence in Figure 6, Region A will come to have higher levels of interpersonal inequality than Region B. This is because Region A concentrates GPT-complementary workers who benefit from rent sharing, as well as the rents themselves that accrue to the entrepreneurs that hire them. Meanwhile, although both regions contain traditional workers, the relative abundance of traditional workers means that the inter-regional wage gradient for these workers will be flatter. Hence, while average wages in A will be higher than B, so will local levels of income inequality.

The outcome of these interlocking dynamics is therefore strong interregional divergence in terms of real and nominal wages, housing prices, second nature amenities, and local interpersonal inequality. Figure 9 depicts these interlocking divergence processes. In Diamond (2016), region A clears on wages and amenities, but region B clears on wages and housing costs, with non-homothetic preferences. It seems improbable that this is a utility-equalizing equilibrium, as in Glaeser (2008). It is admittedly difficult to observing and quantify the value of amenities, experience, and socio-economic connections, but there is increasing shards of evidence that the post-1980 period tends toward a non-equalizing divergence equilibrium (Kemeny and Storper, 2012; Gaubert, 2019; Diamond, 2016).

Figure 10 depicts the reversal of these tendencies. As revolutionary technologies mature, their constituent ideas become more widely understood, standardized and codified. They can increasingly be internalized and imitated by firms in locations far from their creation. Formerly-GPT-complementary firms and workers no longer enjoy strong benefits from co-location or attachment to the initial agglomerations of the industrial revolution in question. Combining these declining benefits with the relatively high costs in Region A, GPT-complementary firms will find it possible relocate to Region B. They adopt locational strategies that are cost-reducing, involving long-distance physical and informational transactions.

Figure 10. The process of convergence



Note: Authors' elaboration

At this point, the arbitraging opportunities shift for both GPT-complementary and traditional workers. The relocation of GPT-complementary firms spreads demand for GPT-complementary labor. The urban wage premium for GPT-complementary workers declines, both as firms can leave the high-cost agglomeration, and also because monopoly rents in these activities are in decline. Wages for GPT-complementary workers therefore begin to converge between Region A and Region B.

Nominal and cost-adjusted income gaps between A and B decline, as do gaps in amenities. The spreading out of GPT-complementary workers equalizes local levels of interpersonal inequality. As traditional workers are relatively abundant, the movement of GPT-complementary workers from A to B will still raise the quality of amenities in B, leading toward the elimination of amenity gaps separating the two locations.

These convergence forces become progressively stronger – albeit at a strong cost to Region A, which now appears to be in relative decline – until a new GPT is introduced. This sets off a new round of divergence.

It is difficult to predict whether the Region A of one industrial revolution will be the Region A of a subsequent revolution (Storper et al., 2015). We have evidence of some types of Region A emerging as high-income centers across different industrial revolutions but also of some prosperous Region A-type city-regions not capturing prosperity in subsequent revolutions. Turbulence in the performance of specific regions thus may be a feature of the wave-like patterns of the economy.

The value of this integrated explanation of divergence and convergence lies not just in its addition of an historical detail to an otherwise convincing identification story for each period, but to getting the identification of spatial equilibria right by understanding the ways in which different elements (notably labor demand and labor supply) of spatial behavior change together, as well as the hierarchy of causes among them. The initially spatially-uneven nature of each revolution generates strong centripetal forces of demand for those workers with skills that are complementary to the revolution. As the revolution matures and its technologies are codified and routinized, their use and their subsequent development diffuse geographically and with them, so do the skilled workers and the returns to skills. One can link these basic changes to consequent shifts in housing markets and amenities, with ramifications for workers' arbitraging opportunities and locational choices, ultimately linked to overall patterns of interregional inequality. This in turn allows us to see that the various dimensions of a spatial equilibrium – migration, housing, amenities, and wages – move together in coherent but inverted patterns in each period.

5. Conclusion: a more robust way of interpreting evidence and identifying policy challenges

The way we frame spatial economic convergence and divergence is important to the practical relevance of spatial economic analysis. If we believe that markets almost always push toward the elimination of interregional income gaps – whether nominal or real – then research and policy attention will be trained on eliminating barriers to such mean reversion. But if markets efficiently push in other periods toward increased interregional inequality, then we require different research questions and policy orientations for these periods. For example, policies aimed at eliminating frictions to mobility are much more likely to achieve their goals in a period in which technology diffusion dominates as compared with a phase in which agglomeration forces dominate (Duranton and Venables, 2018). In divergence periods, policy efforts might instead concentrate on reaping the benefits of polarization, while recognizing that it will generate left-behind regions whose problems cannot be dealt with principally by promoting convergence in the friction-reducing way (Rodriguez-Pose, 2018; Austin et al, 2018). It can also be anticipated that, in divergence periods, there will be intense problems of social and spatial inequality in prosperous regions, challenging existing political institutions. But these are not going to be dealt with by encouraging spatial dispersion (Florida, 2017; Moos et al, 2017; Baum-Snow and Pavan 2013). In other words, without a structural framework that involves historical change and a realistic view of technology and its geography, we will lack the context required to frame research on regional development, and to correctly interpret empirical patterns and relationships.[12]

This perspective also offers a more robust explanation of the obstacles facing today's left-behind places. More than high housing costs, it is the skill-bias of contemporary urban labor demand that deters the outflow of less-skilled workers from left-behind places. Because of the concentration of key skilled activities and further endogenous technological change in those the resulting agglomerations, skill requirements in these superstars tend to be both high and specific. The accumulation of amenities in these locations drives up housing costs, but in a way that suggests that total utility is still higher for such workers in these places than anywhere else (Couture et al, 2019; Diamond, 2016). Concretely, the mid-20th-century image of Midwesterners with high school diplomas moving to the Los Angeles 2020 to take jobs in manufacturing is truly anachronistic (Rodriguez-Pose and Storper, 2019; Anenberg and Kung, 2018). Los Angeles, like most superstar cities of today, is marked by its highly polarized labor market, with a skilled upper tier working in technical, creative and professional services, and a lower-skilled tier populated by immigrants. Where would medium-skilled persons from far-away regions fit in LA? As Autor (2019, p.33) puts it, the fall in migration of the less skilled is less due to a failure of arbitrage as to a fall in the "economic allure of urban labor markets for the non-college educated and a concomitant rise in their allure for the college-educated."

The reference points from the 1940 to 1980 period that figure centrally in many spatial equilibrium models and theories of development are of limited application today, and hence those models are also of limited use in dealing with current challenges. The prospects for using policy for spreading the skilled parts of the economy and their economic benefits will

[12] This does not imply using the past as predictor for any future wave of convergence or divergence. Each industrial revolution unfolds against a backdrop of other conditions – trade costs, demography, institutions, among others – and each revolution has possible different elasticities of skilled labor demand. Whether and exactly how a convergence phase of the third industrial revolution will take place, or whether a Fourth Industrial Revolution will graft itself onto the present divergence phase, and how it will affect interregional inequality, are open questions. The point of the demand-side model argued in this paper is to correctly understand whether the fundamental forces lean toward convergence or divergence at any moment in time.

not be identified with models that were created to understand convergence period. Likewise, the people-based policies that form the basis of our traditional thinking about how to encourage migration are no longer working as they did during the convergence period. The stakes for understanding these relationships are high. As *The Economist* (2016) states, "…if orthodox economics does not come up with an answer, populist insurgents will."

References

Abel, JR, and Deitz, R. (2019). Why are some places so much more unequal than others? *Economic Policy Review* 25,1: 58-75, December, Federal Reserve Board of New York.

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of economic literature*, *40*(1), 7-72.

Aguilar-Retureta, J. (2016). Regional income distribution in Mexico: New long-term evidence, 1895–2010. *Economic History of Developing Regions*, 31(2-3):225–252.

Anenberg, E. and Kung, E. (2018). Can more housing supply solve the affordability crisis? evidence from a neighborhood choice model. *Regional Science and Urban Economics*.

Austin, B., Glaeser, E., and Summers, L. H. (2018). Saving the heartland: Place- based policies in 21st century America. In Brookings Papers on Economic Activity.

Autor, D. (2019). Work of the past, work of the future. *AEA papers and proceedings* 109:1-32.

Autor, D. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–97.

Baldwin, R. (2006) Globalisation: the great unbundling (s). Economic Council of Finland, 20: 5-47.

Barro, R. J., Sala-i-Martin, X., Blanchard, O. J., & Hall, R. E. (1991). Convergence across states and regions. *Brookings papers on economic activity*, 107-182.

Baum-Snow, N; Pavan, R. (2012). Understanding the city size wage gap. *Review of Economic Studies* 79,1: 88-127.

Baum-Snow, N; Pavan, R. (2013). Inequality and city size. *Review of Economics and Statistics.* 95,5: 1535-1548.

Baum-Snow, N; Freedman, W; Pavan, R. (2018). Why has urban inequality increased? *American Economic Review: Applied Economics* 10,4: 1-42.

Beaudry, P., Doms, M., & Lewis, E. (2010). Should the personal computer be considered a technological revolution? Evidence from US metropolitan areas. *Journal of Political Economy*, *118*(5), 988-1036.

Berger, T. and Frey, C.B. (2016). Did the computer revolution shift the fortunes of US cities? Technology shocks and the geography of new jobs. *Regional Science and Urban Economics* 57: 38-45.

Blanchard, O. J., and Katz, (1992). Regional evolutions. *Brookings papers on economic activity*, *1992*(1), 1-75.

Borts, G. H. (1960). The equalization of returns and regional economic growth. *American Economic Review*, 50(3):319–347.

Bresnahan, T. and Trajtenberg, M. (1995). General purpose technologies: engines of growth. *Journal of Econometrics* 65: 83-108.

Browne, L. E. (1989). Shifting regional fortunes: the wheel turns. *New England Economic Review*, (May):27–40.

Broz, J. L., Frieden, J., & Weymouth, S. (2019). Populism in place: the economic geography of the globalization backlash. Mimeo.

Carlino, G. A. (1992). Are regional per capita earnings diverging? *Business Review*, 3.

Carlino, G. A. and Mills, L. (1996). Testing neoclassical convergence in regional incomes and earnings. *Regional Science and Urban Economics*, 26(6):565–590.

Chen, Y., & Rosenthal, S. S. (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics*, *64*(3), 519-537.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *Quarterly Journal of Economics*, 129(4):1553–1623.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of urban economics*, 63(2):723–742.

Connor, D. (2018) Stability and change in the land of opportunity over the twentieth century. Paper presented at the 2018 Annual Meeting of the Population Association of America.

Cooke, T. J. (2013). Internal migration in decline. *Professional Geographer*, 65(4):664–675.

Costa, D. and Kahn, M. (2000). Power couples: changes in the locational choice of the college-educated, 1940-1990. *Quarterly Journal of Economics* 115,4: 1287-1315.

Couture, V., Gaubert, C., Handbury, J., & Hurst, E. (2019). *Income growth and the distributional effects of urban spatial sorting* (No. w26142). National Bureau of Economic Research.

Davis, D.R. and Dingel, J.I. (2019). A spatial knowledge economy. *American Economic Review* 109(1): 153-170.

Davis, M. A. and Ortalo-Magne, F. (2011). Household expenditures, wages, rents. *Review of Economic Dynamics*, 14(2):248–261.

De La Roca, J. and Puga, D. (2017). Learning by working in big cities. *Review of Economic Studies*, 84(1):106–142.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverg- ing location choices by skill: 1980-2000. *American Economic Review*, 106(3):479– 524.

Dorn, D. (2009). Essays on inequality, spatial interaction, and the demand for skills. PhD thesis, University of St. Gallen.

Drennan, M. P., Tobier, E., and Lewis, J. (1996). The interruption of income convergence and income growth in large cities in the 1980s. *Urban Studies*, 33(1):63–82.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In V. Henderson and JF Thisse, eds., *Handbook of regional and urban economics*, Vol. 4, p. 2063-2117. Elsevier.

Duranton, G. and Venables, A. J. (2018). Place-based policies for development. The World Bank.

Economist (2016) Place-based economic policies as a response to populism. *The Economist*, Dec 15th, print edition.

Elsby, M. W., Hobijn, B., & Şahin, A. (2013). The decline of the US labor share. *Brookings Papers on Economic Activity*, *2013*(2), 1-63.

Enflo, K. and Roses, J. R. (2015). Coping with regional inequality in Sweden: structural change, migrations, and policy, 1860–2000. *The Economic History Review*, 68(1):191–217.

Fischel, W. A. (2004). An economic history of zoning and a cure for its exclusionary effects. *Urban Studies*, *41*(2), 317-340.

Florida, R. (2002). Bohemia and economic geography. *Journal of economic geography*, *2*(1), 55-71.

Florida, R. (2017). *The New Urban Crisis.* New York: Basic Books.

Fontagné, L; Santoni, G. (2019). Agglomeration economies and firm labor misallocation. *Journal of Economic Geography* 19,1: 251-272.

Fotheringham, A. S. and Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and planning A*, 23(7):1025–1044.

Ganong, P. and Shoag, D. (2017). Why has regional income convergence in the US declined? *Journal of Urban Economics*, 102:76–90.

Garnick, D. H. (1990). Accounting for regional differences in per capita personal income growth: an update and extension. *Survey of Current Business*, *70*(1), 29-40.

Garrett, M. (1987) *Land use regulation: the impacts of alternative land use rights*. New York: Praeger.

Giannone, E. (2017). Skilled-biased technical change and regional convergence. University of Chicago Working Paper.

Glaeser, E.L. (2008). *Cities, agglomeration and spatial equilibrium.* Oxford: Oxford University Press.

Glaeser, E.L. and Gyourko, J. (2018). The economic implications of housing supply. *Journal of Economic Perspectives*, 32(1):3–30.

Glaeser, E. L., & Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of economic literature*, *47*(4), 983-1028.

Glaeser, E. L. and Maré, D. C. (2001). Cities and skills. *Journal of labor economics*, 19(2):316–342.

Glaeser, E. L., & Resseger, M. G. (2010). The complementarity between cities and skills. *Journal of Regional Science*, *50*(1), 221-244.

Graves, P. (1976). A Reexamination of migration, economic opportunity and the quality of life. *Journal of Regional Science* 16: 107-112.

Greenwood, M. J., & Hunt, G. L. (1989). Jobs versus amenities in the analysis of metropolitan migration. *Journal of Urban Economics*, *25*(1), 1-16.

Gyourko, J., Mayer, C., and Sinai, T. (2013). Superstar cities. *American Economic Journal: Economic Policy*, 5(4):167–99.

Haines, M. R. (2005). Historical, demographic, economic, and social data: the United States, 1790–2002. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research*.

Handbury, J. and Weinstein, D. E. (2014). Goods prices and availability in cities. *Review of Economic Studies*, 82(1):258–296.

Higgins, M. J., Levy, D., and Young, A. T. (2006). Growth and convergence across the United States: Evidence from county-level data. *Review of Economics and Statistics*, 88(4):671–681.

Hsieh, C.-T. and Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39.

Jenkins, S. P., & Van Kerm, P. (2006). Trends in income inequality, pro-poor income growth, and income mobility. *Oxford Economic Papers*, *58*(3), 531-548.

Jenkins, S. P. and Van Kerm, P. (2009), 'dsginideco – Decomposition of inequality change into pro-poor growth and mobility components', v1.0, CEPS/INSTEAD, Differdange, Luxembourg.

Kemeny, T. and Osman, T. (2018). The wider impacts of high-technology employment: Evidence from US cities. *Research Policy*, 47(9):1729–1740.

Kemeny, T. and Storper, M. (2012). The sources of urban development: Wages, housing, and amenity gaps across American cities. *Journal of Regional Science*, 52(1):85–108.

Lee, N., Morris, K., & Kemeny, T. (2018). Immobility and the Brexit vote. *Cambridge Journal of Regions, Economy and Society*, *11*(1), 143-163.

Lee, Y., Lee, B., & Shubho, M. T. H. (2019) Urban Revival by Millennials? Intra-Urban Net Migration Patterns of Young Adults, 1980–2010. *Journal of Regional Science*.

Lin, G.S. (2011). Technological adaptation, cities and new work. *Review of Economics and Statistics,* May, 93 (2): 554-574.

Manduca, R. A. (2019). The contribution of national income inequality to regional economic divergence. *Social Forces*.

Martinez-Galarraga, J., Roses, J. R., and Tirado, D. A. (2015). The long-term patterns of regional income inequality in Spain, 1860–2000. *Regional Studies*, 49(4):502–517.

McLaughlin, G. E., and Robock S.H. (1949). *Why Industry Moves South: a Study of Factors Influencing the Recent Location of Manufacturing Plants In the South*. Committee of the South, National Planning Association.

Milanovic, B. (2019). *Capitalism, Alone: The Future of the System That Rules the World*. Harvard University Press.

Mokyr, J. (1990). Punctuated equilibria and technological progress. *American Economic Review*, *80*(2), 350.

Moos, M., Pfeiffer, D., & Vinodrai, T. (2017) The millenial city: shaped by contradictions. In: Moos, M., Pfeiffer, D., & Vinodrai, T. (Eds.). *The millennial city: Trends, implications, and prospects for urban planning and policy.* Routledge.

Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of econometrics*, 121(1-2):175–212.

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

Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, 5(1):65–103.

Norton, R. D., & Rees, J. (2007). The product cycle and the spatial decentralization of American manufacturing. *Regional Studies*, 41(S1), S61-S71.

O'Neill, D. and Van Kerm, P. (2008). An integrated framework for analysing income convergence. The Manchester School. 76(1): 1-20.

Partridge, M. D., & Rickman, D. S. (2003). The waxing and waning of regional economies: The chicken–egg question of jobs versus people. *Journal of Urban Economics*, *53*(1), 76-97.

Perez, C. (2010). Technological revolutions and techno-economic paradigms. *Cambridge journal of economics*, *34*(1), 185-202

Pomeranz, K. (2000). *The Great Divergence.* Princeton: Princeton University Press.

Roback, J. (1982). Wages, rents and the quality of life. *Journal of Political Economy* 90, 6: 1257-78.

Rodriguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society*, 11(1):189–209.

Rodriguez-Pose, A. and Storper, M. (2019). Housing, urban growth and inequalities: the limits to deregulation and upzoning in reducing economic and spatial inequality. *Urban Studies,* 1-25.

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics*, 74-104.

Rosenberg, N. (1982). *Inside the Black Box: Technology and Economics.* Cambridge, UK: Cambridge University Press.

Rosés, J. and Wolf, N. (2018). Regional economic development in Europe, 1900-2010: a description of the patterns. CEPR Discussion Paper No. DP12749.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., and Sobek, M. (2018). IPUMS USA: Version 8.0 [dataset]. Minneapolis, MN: IPUMS.

Shapiro, J. M. (2006). Smart cities: quality of life, productivity, and the growth effects of human capital. *The review of economics and statistics*, *88*(2), 324-335.

Storper, M., Kemeny, T., Makarem, N., & Osman, T. (2015). *The rise and fall of urban economies: Lessons from San Francisco and Los Angeles*. Stanford University Press.

Storper, M. (2018). Separate worlds? Explaining the current wave of regional economic polarization. *Journal of Economic Geography*, 18(2):247–270.

Storper, M., & Walker, R. (1989). *The Capitalist Imperative*. Oxford: Blackwell.

Schwartzman, F. (2017). Inequality Across and Within US Cities around the Turn of the Twenty-First Century. *Economic Quarterly*, (Q1-Q4), 1-35.

Schumpeter, J. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper & Bros.

Tolbert, C. M. and Sizer, M. (1996). US commuting zones and labor market areas: A 1990 update. United States Department of Agriculture, Staff report.

Appendices

A. Data Appendix

We measure economic disparities largely using wage and salary income. This measure offers advantages over indicators like per capita GDP. One advantage is that, relative to output per head, wages more closely gauge well-being, particularly in more recent years, given observed declines in rent sharing and the overall labor share (i.e. Elsby et al, 2013). Further, we can gather a longer panel for wages, reaching as far back as 1940 using Census microdata, and all the way back to the mid-19th century using aggregate Census information for specific classes of activity, like manufacturing and agriculture.

To characterize the evolution of local labor markets in the post-1930-era, we use individual-level microdata from public-use IPUMS extracts of the Decennial Census and the American Community Survey, spanning 1940 to 2017 (Ruggles et al., 2018). The 1940 Census represents the first in which respondents were asked about their income, hence acts as a backstop before which we cannot use microdata to estimate levels of interregional income inequality. For 1940, 1950 and 1970 we rely on one percent samples of the Decennial, whereas we are able to use five percent samples for 1960, 1980, 1990, and 2000. We use a three-year, three percent sample of the American Community Survey for 2010 (2009-2011) and a one-percent ACS sample for 2017.

From 1940, we use Commuting Zones (CZs) to describe the geography in these data. Commuting Zones are groups of counties that are linked through the intensity of travel patterns, and distinguished by weak inter-area commuting; they therefore effectively represent functionally-integrated economic units (Tolbert and Sizer, 1996). Though this logic makes CZs apt units to explore development patterns in space, no known research has done so over a time span of this length. More generally, CZs have been under-utilized in regional development research. Prior studies have examined Census Regions (Drennan et al., 1996), states (Barro and Sala-i Martin, 1991; Carlino and Mills, 1996), as well as counties (Higgins et al., 2006) and OMB-defined Metropolitan Areas (Giannone, 2017). The former two are overly aggregated, thereby concealing major intra-unit variation. Meanwhile, counties are under-aggregated relative to actual patterns of economic activity in much of the 20th century. Each of these scales therefore suffer from the well-known modifiable areal unit problem (Fotheringham and Wong, 1991), which can profoundly bias results. Metropolitan areas, which are the most common unit of regional analysis today, do not account for portions of the US that are not centered around larger urbanized areas. Moreover, many metros are also incompletely identified in IPUMS samples after 1970. Since unidentified individuals are likely to systematically differ from those identified as a function of their distinct geography, estimates at this scale from available microdata are also vulnerable to measurement error.

Commuting zones are not reported in Census data, hence we must assign individual respondents to them. To do so, we adapt an approach described by Dorn (2009), in which individuals are probabilistically matched to CZs based on the smallest identifiable geography in the Census, which varies across surveys between State Economic Areas, County Groups and PUMAs. We assign each of these basic geographies a probability of belonging to each CZ, based on the population fraction in that CZ. Many locations map directly onto a single CZ. For individuals in locations for which multiple CZs are possible, we replace each observation with a multiple reflecting the number of potential CZ units to which each individual may belong. These receive adjusted person weights that reflect the likelihood that they reside in a given CZ. In other words, individuals are split into components whose size depends on the odds of living in a given CZ based on their recorded basic location. As in Autor and Dorn (2013), we additionally weight individual contributions on the basis of their effective labor supply and also their person-level sampling weight provided by the Census.

We extend Dorn's procedure to 1940, 1960, 2010 and 2017. This gives us 722 contiguous 1990-vintage Commuting Zone units that cover the entirety of the lower 48 states. While CZs will more effectively describe local integration the closer we are to 1990, in earlier years outer edges are likely to contain space empty of people, rather than distinct geographies; we therefore believe the benefits of consistent units are outweighed by the relatively modest bias this assumption is likely to generate. After 1990, multiple CZs might be considered to be part of a larger whole. This problem also affects metro areas (for instance, for both metropolitan areas and CZs, we might argue that in recent years the San Jose and San Francisco areas are best considered to be a single local labor market). We believe our reliance on 1990-era definitions is likely to be problematic only if we project far into the future from 1990.

To capture earlier components of the hypothesized wave pattern, we make use of aggregate information derived from the Census and other sources. Specifically, we use data compiled from earlier population, agriculture and manufacturing censuses, which are available through the Inter-University Consortium for Political and Social Research (ICPSR), and are collectively known as Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 (Haines, 2005). As the title suggests, these data offer a long-run picture of various dimensions of life in the US, with counties as the basic spatial unit. From these sources we obtain county-specific measures of manufacturing employment, capital, and output between 1860 to 1930. On the logic that commuting zones are likely to become progressively less apt as we deviate from 1990, we remain at the county level for this earlier wave of information.

Appendix B: Gini decompositions and bootstrapped SEs

Results in Table 1 and Table B1 depend on decompositions of the change in inequality, as described using the Gini coefficient, proposed and implemented by Jenkins and Van Kerm (2006). They decompose changes in Gini coefficients across two periods into two distinct, additive components, the first capturing distributional mobility, re-ranking or leapfrogging; the second capturing what they describe as the progressivity of income growth, such that we get a negative relationship between initial income and subsequent income growth. In its most general form, the decomposition is as follows:

$$\Delta G(v) = R(v) - P(v) \tag{1}$$

where G represents the generalized Gini coefficient; v is a parameter for which larger values place more weight on income differences between poorer units; R captures leapfrogging; and P measures progressivity. Readers seeking details of the decomposition are referred to Jenkins and Van Kerm (2006).

Whereas Table 1 presents decompositions in terms of percent changes of initial values, Table B1 presents raw estimates, as well as bootstrapped standard errors.

Table B1. Decomposition of changes in interregional income inequality (σ-convergence) into contributions from leapfrogging and β-convergence, US Commuting Zones, 1940-1980.

Year Span	Starting Gini	Ending Gini	ΔGini	Leapfrogging (R)	Progressivity (P)
Panel 1. Major Periods					
1940-1980	15.4	7.0	-8.4	1.9	10.2
	(0.37)	(0.1)	(0.34)	(0.1)	(0.3)
1980-2017	7.0	7.5	0.51	2.5	2.0
	(0.1)	(0.2)	(0.26)	(0.1)	(0.3)
Panel 2. By Decade					
1940-1950	15.4	11.5	-4.0	1.1	5.0
	(0.37)	(0.3)	(0.2)	(0.1)	(0.2)
1950-1960	11.5	11.2	-0.3	1.9	2.1
	(0.3)	(0.2)	(0.2)	(0.1)	(0.2)
1960-1970	11.2	7.7	-3.5	1.0	4.5
	(0.2)	(0.2)	(0.2)	(0.1)	(0.2)
1970-1980	7.7	7.0	-0.6	1.4	2.0
	(0.2)	(0.2)	(0.1)	(0.09)	(0.1)
1980-1990	7.0	7.5	0.4	1.1	0.6
	(0.2)	(0.3)	(0.2)	(0.07)	(0.2)
1990-2000	7.5	7.6	1.2	0.6	0.4
	(0.3)	(0.3)	(0.1)	(0.04)	(0.1)
2000-2010	7.6	7.4	-0.2	0.5	0.7
	(0.3)	(0.3)	(0.09)	(0.04)	(0.09)
2010-2017	7.4	7.5	0.2	1.0	0.8
	(0.3)	(0.3)	(0.1)	(0.07)	(0.1)

Note: N=722 Commuting Zones. Estimates are multiplied by 100 and rounded. Bootstrapped standard errors in parentheses (250 replications). Gini coefficients are conventional in terms of weighting poorer units, i.e. estimated with ν=2. Values are calculated using the Stata program DSGINIDECO (Jenkins and Van Kerm, 2009). For each Commuting Zone, 'income' is defined as estimated average hourly wage and salary income. Incomes are adjusted for inflation to 2015 dollars using Bureau of Labor Statistics CPI. Calculations are weighted by population in the initial period in question. Source data are public-use (IPUMS) extracts of the Decennial Censuses and American Community Survey. Data details are found in Appendix A