The fall and rise of interregional inequality: Explaining shifts from convergence to divergence

Tom Kemeny
School of Business and Management
Queen Mary, University of London
Mile End Road, London E1 4NS
t.kemeny@qmul.ac.uk

Michael Storper
Luskin School of Public Affairs
University of California Los Angeles
3250 Public Affairs Building, Box 951666
Los Angeles, CA 90095
storper@ucla.edu

Please cite as:
Kemeny, T., & Storper, M. (2020). The fall and rise of interregional inequality: explaining shifts from convergence to divergence. *Scienze Regionali*, *19*(2), 175-198. DOI: 10.14650/97084

Abstract

This paper presents new evidence on long-run patterns of interregional inequality in the United States. The evidence points to a fundamental shift from convergence to divergence, prompting a reconsideration of theory. The paper augments existing demand-side theories, linking them to historical perspectives on disruptive innovation and industrial revolutions. It offers an integrated account of convergence and divergence that matches the facts. The resulting perspective directs attention to workers that are complementary to the current revolution; as well as to the contingency of policy effects on whether or not the regions are in a mode of convergence or divergence.

Keywords: regional development; inequality; technological change

JEL codes: R12; O18; O33

1. Explaining the increase in interregional inequality

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 Italy, between the Mezzogiorno and prosperous northern cities like Milan and Bologna. 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 what seems to be a cumulative process, superstar cities are differentiating themselves from the rest of the countries in which they are embedded.

A vibrant debate is underway about this polarization, aiming to understand its causes, consequences, and possible policy responses. The discussion of causal mechanisms is split into two primary narratives: one emphasizing labor supply, the other, labor demand. Those centered on labor supply explain rising inequality as a consequence of worker locational choices, with those choices strongly conditioned by preferences and constraints that vary according to one's level of educational attainment. The demand-side explanations stress the geographical consequences of a skill-biased technological change.

This paper describes long-run patterns of interregional inequality in the U.S., and uses this evidence to deepen demand-side explanation. We describe alternating phases of convergence and divergence in the USA since 1940 at the level of Commuting Zones, relating our findings to the existing empirical literature (i.e. Moretti, 2012; Gyourko, 2013; Giannone, 2017). The view from the long run necessitates changes in the conceptualization of geographical forms of inequality. Specifically, the post-1980 technology shock is an instance of a more general, recurring phenomenon – an industrial revolution – which we argue is the direct cause of episodic periods of rising and falling inequality. Convergence and divergence are both principally driven by the behavior of a specific subset of workers: those whose skills are most directly complementary to the revolution at hand. Understanding the role of the skilled and the geography of the firms that demand them is essential in developing models that capture the interrelated dimensions – wages, migration, housing and amenities – of each phase, and their coherent pattern of interrelated

differences. Only by having such a comprehensive view of causes and effects for each phase – can applied research identify the specific policy challenges of divergence and convergence phases, and avoid applying yesterday's solutions to today's problems.

2. From convergence to divergence in the U.S: The evidence

Though there is a growing empirical literature on interregional inequality in the U.S., a consistent set of data on its timing and dimensions at an appropriate geographical scale. Using data on Commuting Zones (CZs), in this section we highlight four key dimensions of the evolution of interregional economic divergence: (a) changes in overall inequality, or σ-convergence ; (b) changes in the location of college graduates; (c) changes in wage inequality for workers at specific levels of educational attainment; and (d) real wage inequality, measured by adjusting household nominal incomes for housing costs.

Commuting zones are determined according to the intensity of travel patterns and distinguished by weak inter-area commuting (Tolbert and Sizer, 1996).[1] 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 our use of the data are found in the Data Appendix.

(a) Nominal interregional income convergence ends around 1980

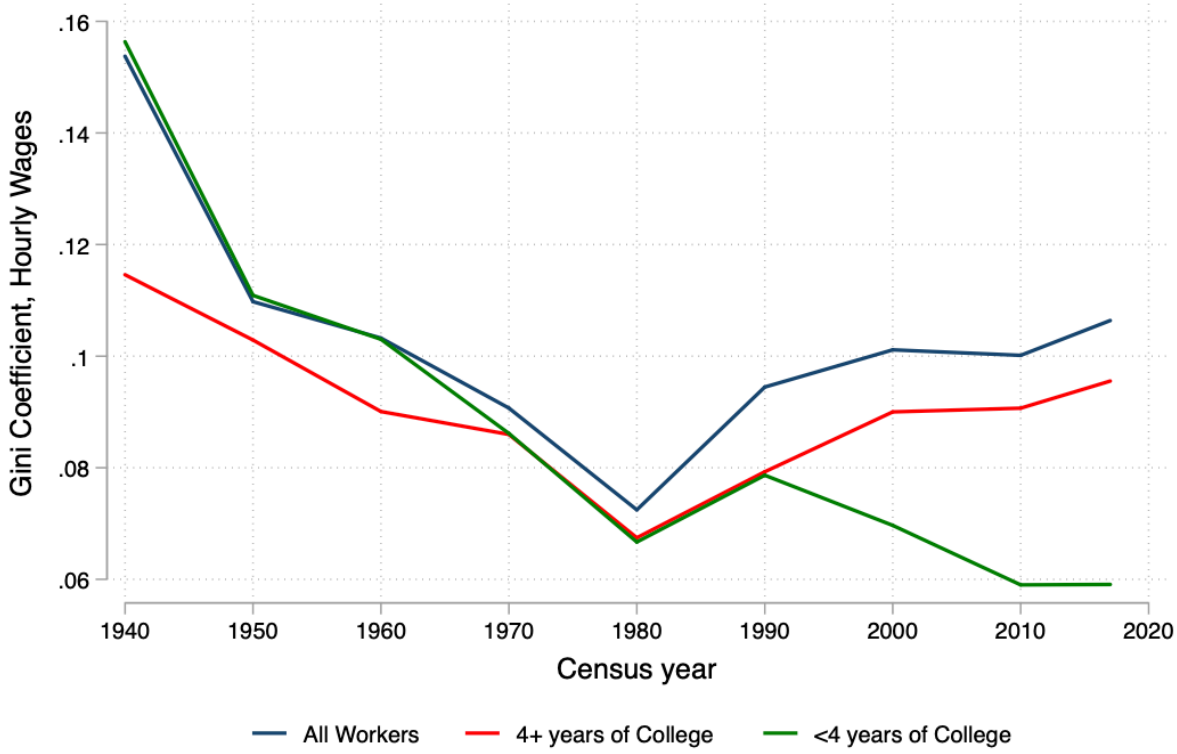
Figure 1 visualizes Gini coefficients of average hourly wage and salary income, tracing the evolution of interregional inequality between 1940 and 2017. Focusing on the blue line representing estimates for all in-sample workers, the figure shows that, from 1940 until 1980, regional differences in incomes were in decline. From 1980 forward, gaps begin to grow if we weight units by population (as we do in the figure); if we do not differentiate by population, interregional income gaps stop declining. This distinction that suggests that large metropolitan areas play a considerable role in the recent turn to divergence.[2][3]

[1] Throughout this paper, we will use terms like 'Commuting Zone', 'region', 'city', 'local' and 'urban' more or less synonymously to refer to economically integrated locations.

[2] In this section we mainly weight by population. We do so because the distribution of CZ populations is broad, ranging in 2017 between a low of 936 (the region centered on Murdo, SD) to a high of 18.8 million (Los Angeles, CA). An increase in average income in Murdo means something very different for the US urban system than an analogous change in LA. If we were to not weight by population, the convergence phase remains largely unchanged, while instead of a shift to divergence after 1980, we get a neutral pattern, with no evidence of either convergence or divergence.

[3] Patterns described in the text are robust to a range of methodological choices. We get strongly comparable results if we measure dispersion using the coefficient of variation of log wages. Results also remain materially similar when we measure incomes using annual rather than estimated hourly wages, and if the sample is restricted to employed

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 the interruption of convergence soon after it began (i.e., Browne, 1989; Garnick, 1990; Drennan et al., 1996). Though Carlino (1992) imagined this to be a brief aberration in the secular drive towards convergence, today we know it has continued. Over 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). Much of this work is focused on the recent divergence, with considerably less attention to the convergence period that preceded it. However, the evidence from the long-run suggests that a robust theory

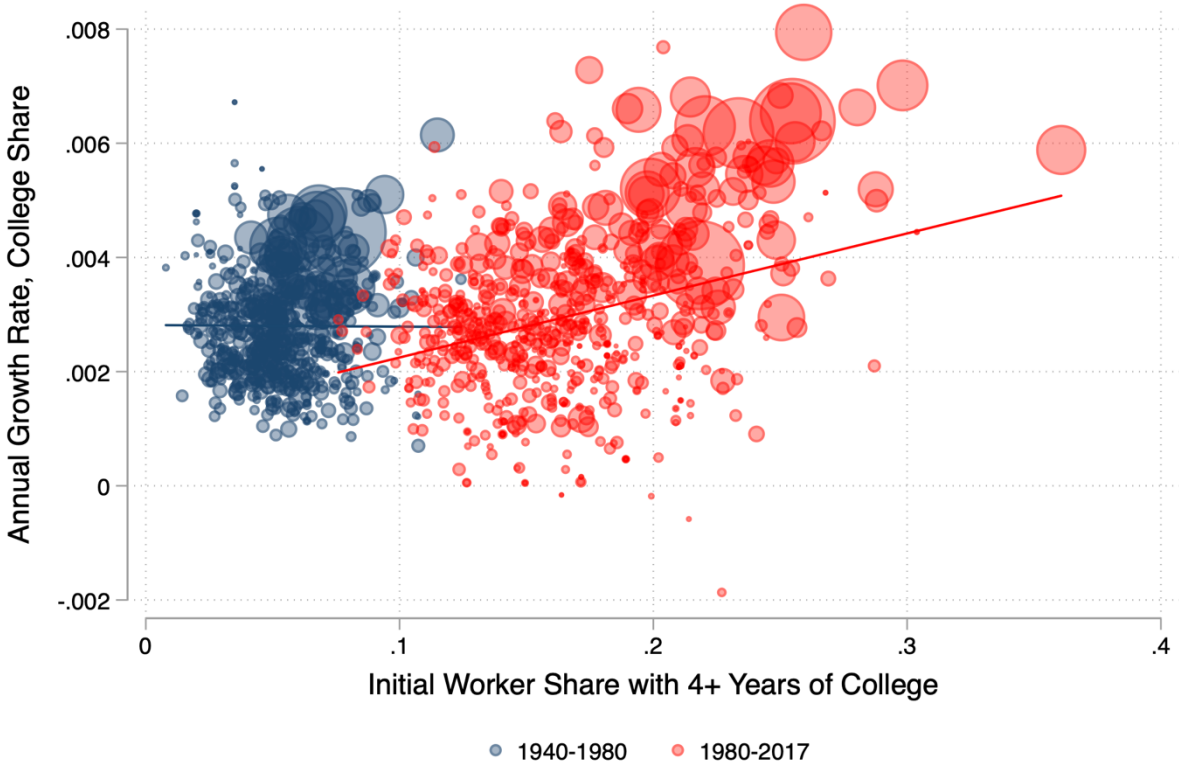
individuals who work at least 35 or (separately) 40 hours per week. Responding to possibility that results could be driven by the one-time shock of the incorporation of the South, we also find consistent results when we omit CZs in South Atlantic, East South Central, and West South Central Census Regions.

needs to explain both convergence and divergence, as well as that which spurs the alternation from one pattern to another.

(b) College graduates have concentrated in space since 1980, but were not doing so during 1940-1980

Most of the literature on growing income inequality in general attributes 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. 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.

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 additionally 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.

And yet 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.

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

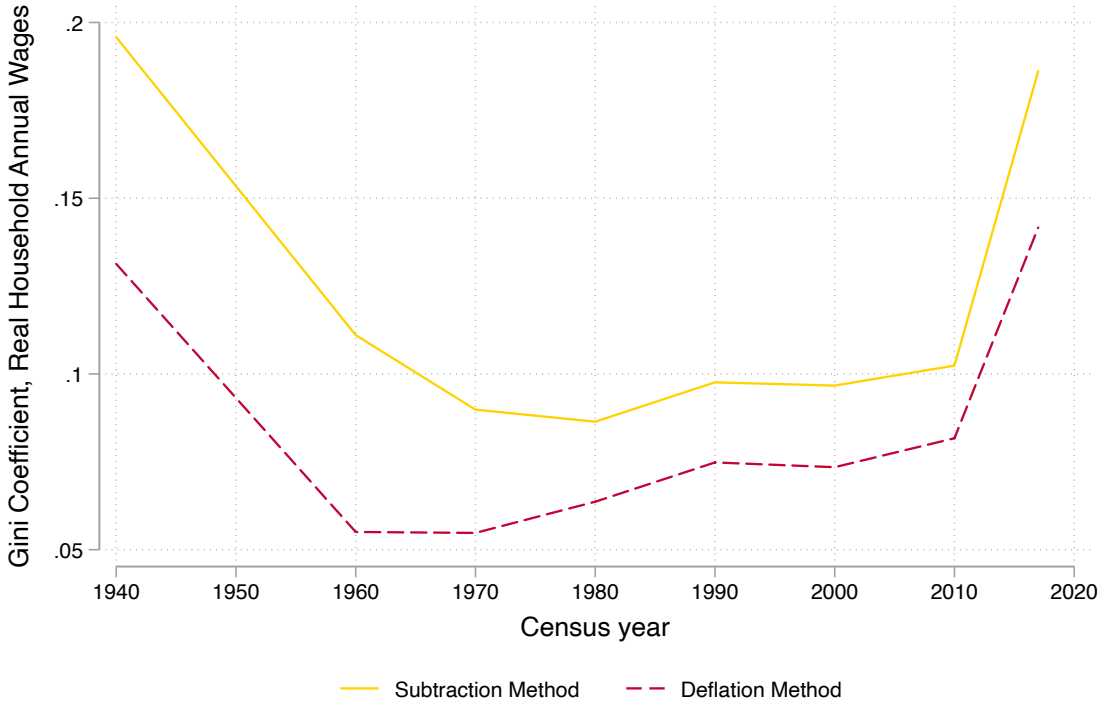
Considering the location-specific returns to skill, we return to Figure 1. The red line traces the evolution of the Gini index for hourly pay across CZs for college graduates only, while the green line captures the pattern for those with less than 4 years of college. These lines tell quite different 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 locationally stratified: 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. The geographical specificity of wages for these workers has diminished fairly consistently over the nearly 80-year study period.

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

While recent income growth has been in more populous and skill-abundant regions, high local prices, and in particular housing costs could negate higher salaries for the college educated, and seriously reduce real incomes for the less educated (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).

General equilibrium effects on prices mean that places that are desirable due to demand and/or appealing amenities ought to have higher housing costs, assuming a less than perfectly elastic housing supply. Even ruling out any market-clearing spatial equilibrium, to compare well-being across places local prices must be taken into account, as is routinely done in an international context using purchasing power parity indices.

Figure 3. Evolution of 'real' interregional income inequality (σ-convergence), subtraction and deflation methods, 1940—2017



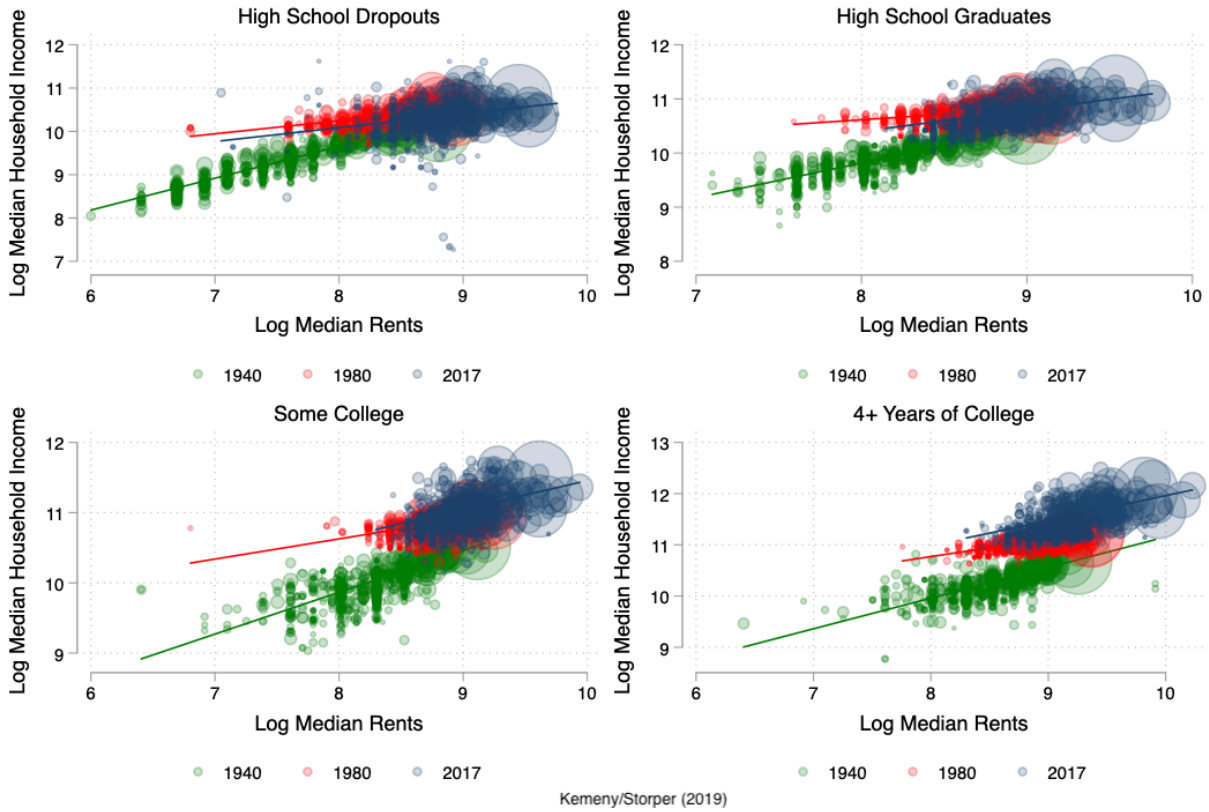
Note: N=722 Commuting Zones. Based on year-specific Gini coefficients estimated using average household real annual hourly wage and salary income for all in-sample households in 1990-vintage CZs, weighted by population. Source data are public use extracts of Decennial Censuses and the American Community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

To enable this kind of comparison, we report the evolution of 'real' (cost-adjusted) household incomes in Figure 3. 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 pursue two methods. First, in order 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 contract 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.[4] 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.

Figure 3 suggests that the long-run evolution of interregional wage inequality does not strongly depend on accounting for housing costs. 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 3 resemble one another. Quantifying this resemblance, the correlation 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. 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.

[4] 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.

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 in empirical terms, nominal income growth in some of these cities – notably places like Boston and Bay Area locations – 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.

It is conceivable that the exclusion of non-housing costs might materially change the results. However, extant research suggests that it is unlikely to reduce the observed advantages of large superstars. 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. Contributions that take a wider view of second-nature amenities suggest that their nature and spatial distribution 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). This suggests that high-income locations will be well endowed with amenities, a correlation confirmed by Kemeny and Storper (2012). It is also consistent with Couture et al (2019), who 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).

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 4 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. The figure shows that 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. This suggests a real inequality dynamic on an interpersonal level, layered within the broader patterns of interregional inequality.

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

Traditional regional economics is marked by explicit or implicit 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). 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). Household locational arbitraging is said to generate the tendency toward convergence of real wages or, in a more expansive version, utility. These are supply-side models with preferences for wages, rents and amenities driving locational choice. In that empirical context, the classical assumption of complete nominal long-term income convergence could be relaxed. A diverse range of preferences were arbitraged through spatial sorting, yielding convergence of a different sort. Rather than nominal incomes, it would be real incomes or utility that, for comparable workers at the margin, ought to equalize. Concretely, an increasing proportion of workers chose 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, and regional gaps in the composition of skilled and unskilled were declining.

A different set of preferences and constraints 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). There is strong empirical support for the idea that skilled workers gain from working in cities (Davis and Dingel, 2018). College graduates make each other more productive when they are co-located (Moretti, 2004; Shapiro, 2006). 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 gains in terms of productivity and wages that graduates have enjoyed in superstar cities. A second possibility is that graduates are making locational choices to maximize on amenities. Specifically, they 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 that do not resemble those of their parents.

To explain the recent sorting behavior of (especially native) non-graduates and the fates of left-behind places, supply-side arguments emphasize barriers to migration. The primary argument centers on housing. 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 primary assumption behind this prediction is that left-behind regions contain deep pools of workers who would actually prefer to be in superstars; in their simulations of an urban system with lower housing costs, major increases in migration from less prosperous to more prosperous regions would take place. Today's most populous cities become dramatically more populous, 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.[5] 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 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 neither 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).

The principal alternative view is one that starts from the labor demand as the direct determinant of interregional inequality. Behind demand lies technological change. In its simplest form, the argument is that technology has reshaped the nature and geography of firms' demand for workers. Computers and related new technologies made graduates more productive, while they replaced a great deal of routine middle-skill work. 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 workers can find suitable work has been reshuffled.

[5] 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.

A wealth of theory and empirical work supports these ideas. On the more empirical front, 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) suggest 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 locational differences in terms of the composition of output mean that cities are rewarded differently. Locational advantages then become self-reinforcing, partly through the endogenous provision of amenities.

But others argue that an average we cannot derive all the divergence from the pre-1980 regional factor endowments. Thus, Baum-Snow and Pavan (2012, 2013) and Baum-Snow et al (2018) argue that technological change alone is not enough to account for the geography of capital-skill complementarities that have come about since 1980. The technology shock doesn't just reinforce the geography of 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 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. It 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).

There are some indirect hints in the literature about the historical roots of today's inequality. Milanovic (2019) argues convincingly that that top incomes today involve a greater overlap between labor income and capital income than in previous periods (such as the early 20th century), thus the key role of the skilled. 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 is very little literature that aims to capture both convergence and divergence in a single framework. To do so, requires an explanation of alternations in labor demand and its geography. This type of account would consider how major, disruptive technological change of the kind commonly described as an industrial revolution could fundamentally reshape the nature and geography of demand for workers. The initially spatially-uneven nature of the revolution could generate strong centripetal forces of demand for those workers with skills that are complementary to the emergent technological regime; the eventual codification of its underlying ideas would spur a subsequent diffusion. 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.

Think about 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. 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. 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 is well underway, with relative deindustrialization of many of the industrial cities in the Northeast and Midwest compared to the 1880 to 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 reshaped the nature and geography of the economy. 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 others. The labor requirements of these new activities are skill-biased, and their development is initially strongly spatially concentrated.

These narratives suggest an opportunity to link current research aiming to identify the causes of today's divergence, and well-established perspectives linking technological change and development in the fields of economic history and economic geography. 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). Whether it is sorting or matching, and whether or not these behaviors are shocked by skill-biased technological change and augmented by endogenous amenities, we still need to account for the shift from convergence to divergence.

The historical literature on technological change offers this missing element. It suggests that, behind any significant changes in sorting or matching behaviour of firms, is the degree of novelty of fundamental technologies and the frontier. 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, 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.

In turn, technological change exhibits some regularity, consisting of waves of disruptive innovation (Schumpter, 1942; Bresnahan and Trajtenberg, 1995; Mokyr, 1990; Perez, 2010; Pomeranz, 2000). Weak codification or high ongoing potential for extension and recombination of technologies leads to high dependence on frequent and complex interactions among inventors and innovators, on the one hand, and supply chains designed to incorporate the possibility of rapid change, on the other (Rosenberg, 1982). These fundamentals have been incorporated in models of the geography of disruption, consisting of a cycle of geographical concentration followed by de-agglomeration or diffusion (Pred and Hagerstrand, 1967; Storper and Walker, 1989). When major new technologies arise, they tend to be unevenly distributed in geographical space and then to consolidate a geography of skilled agglomerations; as they age, they disperse or diffuse. This seems to be true of the key skilled labor that is complementary to each technological revolution. 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.

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 suggests the possibility of a previous round of divergence in the late 19th and early 20th century – in the early part of the second industrial revolution – that has not yet been explored in the economic geography literature. In any case, the 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 rise in such behaviors after a long period of decline, due to an industrial revolution.

Such a 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 limits 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). In contrast to the simulation in Moretti and Hsieh (2019), given a reality in which the wages for the unskilled have converged over space, it is improbable under any scenario that labor from left-behind regions that has not already migrated would respond to anything other than drastic shock to the cost of housing. 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 Midwesterners fit? 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 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. Understanding the inner logic of technology disruption, agglomeration and diffusion underscores the key role of skilled complementary labor in translating the geography of technology into the geography of incomes. This in turn allows us to see that the various dimensions of spatial equilibrium – migration, housing, amenities, and wages – move together in coherent but inverted patterns in each period. The gain is not just in historical perspective, but in fundamental explanatory power.

4. Conclusion: for each period, distinctive causes, consequences and policy challenges

In terms of basic knowledge and policy implications, the way we frame spatial economic convergence and divergence is of deep importance.

If we believe that markets predominantly 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, eliminating frictions to mobility in a period in which technology diffusion dominates are likely to have success, but they are likely to be much less effective in a phase in which agglomeration forces dominate (Duranton and Venables, 2018). Indeed, 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 results.[6]

That is nowhere more important than in today's context, where there are great challenges to reviving left-behind regions, as well as to managing the inequalities within superstar cities. The

[6] 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.

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 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. Thus, as *The Economist* (2016) states in referring to the current inter-regional divergence, "…if orthodox economics does not come up with an answer, populist insurgents will."

References

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.

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.

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.

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.

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, 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.

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.

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.