The changing shape of spatial income disparities in the United States

Tom Kemeny[†] and Michael Storper[††]

[†]University of Toronto
[††]University of California, Los Angeles & London School of Economics

Abstract

Spatial income disparities have increased in the United States since 1980, a pattern linked to major social, economic and political challenges. Yet, today's spatial inequality, and how it relates to the past, remains insufficiently well understood. The primary contribution of this article is to demonstrate a deep polarization in the American spatial system – yet one whose character differs from that commonly reported on in the literature. The increase in spatial inequality since 1980 is almost entirely driven by a small number of populous, economically-important, and resiliently high-income 'superstar' city-regions. But, we also show that the rest of the system exhibits a long-run pattern of income convergence over the study period. A secondary contribution is historical: today's superstars have sat durably atop the urban hierarchy since at least 1940. Third, we describe six distinctive pathways of development that regions follow between 1940 and 2019, with certain locations catching up, falling behind, and surging ahead. We explore the role played by initial endowments in driving locations down these pathways, finding population, education, industrial structure and immigrant attraction to be key distinguishing features. These insights are enabled by a fourth contribution: methodologically, we use group-based trajectory modeling – an approach new to the field that integrates top-down and bottom-up views of the evolving national spatial system. We conclude by exploring implications for the mid-21st century.

Keywords: inequality, geography, cities, convergence, economic history

1 Introduction

The classic finding on spatial income disparities in the United States comes from Barro and Sala-i-Martin (1991): between 1880 and 1988, poorer states exhibited higher income growth rates than richer ones. That century of inter-state income convergence was long interpreted as a sign of a successful long-term American integration experience, as well as one that confirms the assumptions of convergence-oriented growth theories (Barro and Sala-i-Martin, 1992).

Yet, since around 1980, spatial income disparities in the United States have either increased (Carlino, 1992; Manduca, 2019; Kemeny and Storper, 2020; Gaubert et al., 2021; Moretti, 2012) or stopped declining (Ganong and Shoag, 2017). Spatial inequality is of both theoretical and practical interest. A growing body of evidence demonstrates that communities and regions profoundly shape the well being of the individuals and households who live in them. Geographical differences in incomes have been linked to disparities in long-term non-employment (Austin et al., 2018); career ladders (De La Roca and Puga, 2017; Eckert et al., 2022); intergenerational social mobility (Chetty et al., 2014; Connor and Storper, 2020); health (Singh et al., 2017; Case and Deaton, 2020); race-based exclusion (Sitaraman et al., 2020); as well as cultural and political polarization (Cramer, 2016; Rodríguez-Pose, 2018). Place-based income inequality is all the more worrying since, in recent decades, internal migration – an important mechanism by which disadvantaged people escape conditions that limit their achievement, and match themselves to opportunities – has also been in decline (Molloy et al., 2011).

Despite the theoretical and practical significance of this topic, our understanding of American spatial income inequality remains incomplete. Most analyses are based on indices that summarize aggregate between-place inequality, capturing overall trends but revealing little about the changing fortunes of the many and diverse places within the system. At the opposite extreme is a 'bottom-up' literature of case studies that trace the shifting fortunes of specific cities, such as Boston (Glaeser, 2005) and Cleveland (Lamoreaux et al., 2007). While rich, such studies cannot reveal the extent to which chosen cases represent experiences shared across places, nor do they capture how individual trajectories come together in a broad pattern of divergence or convergence.

This article seeks to address these gaps. Using tools that integrate top-down and bottom-up perspectives on place-based development, we shed new light on changing geographical income disparities in the U.S. between 1940 and 2019.[1] We analyze shifting aggregate patterns in incomes across the urban hierarchy, before presenting the results of a group-based trajectory model (GBTM) that reveals how convergence and divergence are underpinned by distinctive, coherent place-based trajectories. GBTM, a form of unsupervised machine learning, has been used widely in fields including psychology, medicine and criminology to derive latent groups based on individual patterns of change (i.e. Eggleston et al., 2004; Colen et al., 2018; Neil et al., 2021). To our knowledge, such methods have not yet been applied to questions in economic geography.

[1] There is no perfect single measure that captures all the relevant inequalities of places within a spatial-urban system; in this research, we rely on average income as a robust, if not perfect, indicator of the underlying quality of the economic development that is occurring in a place, and the distribution of such quality of regional development across the system. See Appendix C for wider discussion of this topic.

Deploying this approach leads to original insights. First, we demonstrate that since 1980, the aggregate expansion of spatial income gaps masks a bifurcated pattern of development. We observe considerable convergence among the majority of commuting zones. At the same time, we find that overall divergence is driven by a modest number 'superstar' city-regions, whose influence on the system's divergence comes both from significant growth in their average income levels as well as their considerable share of economy-wide population. Second, we show that today's superstar cities have in fact remained atop the spatial income hierarchy throughout the 79-year study period, but that the extent of their divergence from the rest of the system has waxed and waned. Third, when we consider local economic change beyond the superstars, we observe steady convergence toward a rising all-region mean. Fourth, unlike in some standard theories, convergence is not simply a win-win process. Some places are experiencing relative declines toward the mean, while others are rising up from initially low income levels. What convergence there is, therefore, signals a range of experiences: transformative catch up for some regions, while for others a notable relative slippage. These complex pathways cannot easily be subsumed under the popular binary of 'superstar' locations (Gyourko et al., 2013; Galbraith and Hale, 2014; Diamond, 2016) and 'left-behind' or 'excluded' places (Wuthnow, 2019; Spicer, 2018). Finally, we explore the role of initial economic, social and demographic features in selecting regions into trajectories, opening up avenues for future work to explain why certain places have followed particular pathways of development.

2 Literature: Perspectives on Spatial Income Disparities

An understanding of spatial income disparities is important because of the central role they play in theoretical and empirical research in economic geography, as well as the economics of cities, labor, and growth. For instance, the recent rise in geographical inequality appears to challenge the conventional wisdom that, through the mobility of people and capital, lower-income regions should catch up to richer ones (Solow, 1956; Glaeser, 2008). In studies of the post-war convergence period, the massive redistribution of the American population toward the South and West was interpreted as confirmation that high mobility and mean reversion worked hand in hand. Yet, since around 1980, internal migration has sharply declined (Molloy et al., 2011). There is little consensus on the causes of the migration slowdown. Some, such as Hsieh and Moretti (2019) and Ganong and Shoag (2017) suggest it is the result of bad policies, in particular excessive land use regulation. Others argue for a structural view of the slowdown, positing that high internal rates of migration in the past are best understood as an historical one-off, rooted in a phase of American land abundance and expansion that has now been exhausted (Austin et al., 2018; Grandin, 2019). At the same time, there is wide agreement that the mobility that remains in the contemporary period is highly spatially selective by level of education, and also perhaps in terms of ability, and that these features are different from the 1940-1980 period. The additional observation that amenities have piled up in the same places where nominal incomes have grown the most (Couture et al., 2019) suggests that, even if cities with high average income levels have become increasingly expensive places to live, they likely offer higher average utility levels.

Recognizing that the system today appears distinctly non-convergent, efforts have been redirected towards identifying possible causes. Research has mostly focused on explaining the observed concentration of highly-educated, high-wage workers in selected cities, aiming to both identify its income effects (Baum-Snow and Pavan, 2012, 2013) and explain its determinants (Diamond, 2016; Autor, 2019; Card et al., 2021; Dauth et al., 2022). On the causal side, there are are four broad lines of inquiry. One literature emphasizes urbanization. City size or density may influence development by spurring specialization and diversity, attracting and keeping skilled workers, and offering them opportunities to learn from one another (Ciccone and Hall, 1996; De La Roca and Puga, 2017). More generally, this work considers that territorial development may be becoming more structurally urban, with the implication that an important source of today's divergence may be the growing scale of agglomeration. A second line of research is on the local supply of skilled or suitably-educated workers. Some of this work stresses assortative matching among skilled workers, with key advantages stemming from the networking, knowledge spillovers, or experience effects of co-locating with other skilled people, as well as shared preferences for similar amenities (Glaeser and Maré, 2001; Moretti, 2004; Autor, 2019). A third perspective highlights the role played by industrial structure in shaping labor demand and thus long-term development (i.e., Chinitz, 1961; Kim, 1998; Kemeny and Storper, 2015). Work of this kind identifies key sectors, and often behind those technologies, linked to particular historical epochs. For instance, Lamoreaux et al. (2007) considers that Cleveland's decline in the second half of the 20th century can be explained by its failed transition away from manufacturing. A fourth perspective considers the developmental effects of social and leadership structures. Ingrained and network-based beliefs and conventions differently focus the attention of firms, educational institutions, and other key actors (North, 1987; Acemoglu et al., 2005). Deep-rooted regional differences in these institutions shape (and are shaped by) divergent economic outcomes (Putnam et al., 1992; Storper et al., 2015; Rodríguez-Pose, 2020; Sitaraman et al., 2020; Petach, 2021).

There remain important gaps in these bodies of research. Though they consider sources of differential regional performance, we know little about how different fates are situated in the context of overall reshaping of an urban-regional system. Consider that when a system is trending toward overall income convergence, it might be a 'win-win' way for all places, through different income growth rates; or it could take a 'win-lose' form, through downward income change for countries or regions at the top (Samuelson, 2004). Even if downward change is purely relative, its meaning in a location that is in the process of losing its former dynamism will differ sharply from a place experiencing absolute and relative catch up. Divergence, too, may reflect many different possible combinations of income development for different regions. These possibilities are not fully captured by either β or σ convergence (O'Neill and Van Kerm, 2008). Such considerations suggest that, within a shifting income distribution, there may be unobserved and yet distinct groups of regions identifiable through their trajectories; an idea that resembles conditional or club convergence in the international development literature (Durlauf and Johnson, 1995; Galor, 1996). Thus far, at the regional scale, descriptions of latent or conditional development clubs have been limited, anecdotal, and largely concerned with cross-sectional income gaps rather than dynamics of inequality.[2] Considering contemporary regional inequality, most studies are limited to a binary view of such clubs, consisting of high-performing 'superstar' cities (Gyourko et al., 2013; Galbraith and Hale, 2014; Diamond, 2016) and at the opposite extreme, 'left-behind' or 'excluded' places (Wuthnow, 2019; Spicer, 2018).

Also unexplored in such accounts is how groups of regions that may exist in the present – whether superstars, left-behind or some other state of being – relate to the past. Top-down, aggregate convergence trends and cross-sectional groupings do not capture the dynamic properties of such clubs. They are also silent on causal questions around why certain types of places rise and then decline, others appear to be resilient in overcoming challenges, others get locked into apparently durable states of stagnation, and still others manage to break out of poverty and move upward within the system. Such issues of resilience, persistence and turbulence have been the focus of a largely distinct literature aiming to understand relatively short-run differences in regional responses to shocks, like the financial crisis that began in 2008 (Christopherson et al., 2010; Davies, 2011; Martin, 2011), or the COVID-19 pandemic (Gong et al., 2020). Researchers in this area have been interested in identifying properties of resilient economies, notably around the supply of skilled labor and industrial structure (Clark and Bailey, 2018). Chapple and Lester (2010), for example, find that resilient U.S. metropolitan areas retain their manufacturing sectors, attract immigrants, and are relatively innovative. Another attempt to explain pathways of regional economic change comes from evolutionary economic geography. Neffke et al. (2011) and Boschma (2015), argue that regional fortunes depend on the extent to which a region's firms are able to adopt new outputs and processes that are closely related to existing areas of technological strength. But such work leaves unexplored the implications of the sum total of such evolutions in terms of convergence or divergence in the urban-regional system as a whole.

Case studies on these topics offer rich, location-specific narratives, while generating hypotheses about drivers of longer-run success that can be more systematically tested. Glaeser (2005), for instance, argues that the skills of its workforce enabled Boston to adapt to changing circumstances, remaining prosperous over a 400-year period. But studies like these are silent about the generality of the individuals pathways they sketch. Meanwhile, large-sample statistical approaches leave unexplored the connections between individual cases and wider dynamics of spatial inequality. This points to a gap in research that integrates both systemic and place-oriented perspectives. Such work would be fully open as to the nature and shape of regional economic development trajectories, whether they be persistent, resilient, path-dependent or transformative. Further, such pathways ought to be defined in part by the relationship of every region to the distribution of incomes across the urban system as a whole. In this paper, we take up this challenge, by exploring the relationship between system-wide patterns of spatial income inequality, and the distinctive, long-run pathways of development that constitute them.

[2]Though framed differently, recent work by Connor et al. (2022, 2023) and Houlden et al. (2022) can be thought of as exceptions.

3 Empirical methods

We address the issues raised in the review through three related inquiries. First, we report on overall trends in spatial income disparities over the period 1940 to 2019. To do so, we rely on measures of dispersion, chiefly the Gini coefficient, as well as changes in the ratio of local average incomes to the all-region ('national') average. Second, we consider the degree of persistence in the distribution over time, seeking to understand whether higher and lower performing locations remain entrenched in their relative positions over the 79-year study period. To explore this, we describe correlations of ranks over time, as well as transition matrices across quantiles of the interregional income distribution. Third, we investigate the premise that, while locations' experiences of long-run income growth has not been one-size-fits-all, there exist commonalities – as yet unobserved – tying subsets of places together. Our aim is to identify these latent groups of regional economies. We assume that membership in a group need not be reducible to a single preexisting characteristic, such as initial income or education. To identify groups of cities that evolve in coherent ways across the study period, we use a statistical approach known as group-based trajectory modeling (GBTM). In estimation of the trajectory models, we also consider the role of key hypothesized drivers of regional development, resilience and growth in selecting particular regions into a given trajectory.

GBTM is a form of unsupervised machine learning whose purpose is to identify internally-homogeneous latent groups on the basis of common time paths on a particular outcome. In our case, we are interested in detecting groups of regions on the basis of the evolution of each region's average income level relative to the all-region average. When group membership is defined by changes in an outcome over time, GBTM offers tangible advantages over competing approaches. Specifically, unlike groups formed using ad hoc or *a priori* distributional features, GBTM offers a formal statistical model in which the existence and number of groups emerges out of the data itself (Nagin, 2005). Moreover, rather than simply assigning units to classes, assignment is probabilistic, accompanied by standard errors. Compared to cluster analysis, GBTM is both explicitly dynamic and built upon standard maximum likelihood estimation; as such it benefits from being consistent and asymptotically normally distributed (Greene, 1990). Further, unlike k-means clustering, the analyst can adjudicate between competing GBTM models using standard goodness-of-fit statistics. GBTM approaches also share commonalities with latent growth curve analysis and hierarchical models, in that each aims to measure potential heterogeneity in a population over time. The fundamental difference between these and GBTM is that the former two assume a continuous distribution across units that can be modeled using a multivariate normal distribution of parameters; whereas GBTM assumes variation can be grouped into distinctive categories (Nagin, 2005). To underline the point: growth curve models are apt when the analyst expects a common underlying developmental process with differences in rates expected to be a function of included predictors. With the methods used in the present paper, we do not expect units to follow a common developmental process. Rather, our interest lies in exploring the idea that groups of regions may follow distinctive pathways of economic development, with distinctive causes. We therefore aim to approximate such groups, identify the units that are likely to belong to each group, and to explore factors associated with belonging to a given group.

Applying this approach to our context, for commuting zone i in time T, let the longitudinal sequence of observed measurements on incomes, or more specifically local average wages detrended against the mean across all locations be:

$$Y_i = \{y_{i1}, y_{i2}, ..., y_{iT}\} \tag{1}$$

Let $P(Y_i)$ be the probability of Y_i, to be estimated using maximum likelihood, using the censored normal likelihood (or tobit) function to derive estimates. The aim in estimation is to determine the set of parameters, Ω, that maximize $P(Y_i)$. These parameters – polynomial functions of time – are assumed to be unique to each trajectory, defining their shape, as well as the probability that a given unit belongs to a group j, the collection of which, J, describes the finite count of discrete groups. Given that group membership is not observed, but rather that determining membership is a primary aim in estimation, to estimate $P(Y_i)$ requires the summing of conditional likelihood functions $P^j(Y_i)$. The underlying finite mixture model is given as follows:

$$P(Y_i) = \sum_{j}^{J} \pi_j P^j(Y_i) \tag{2}$$

where $P(Y_i)$ is the unconditional probability of observing commuting zone i's sequence of time-varying measurements of the dependent variable, Y_i; π_j is the probability that a given observed unit belongs to group j. This is called a mixture model because it is assumed that the population is composed of a mixture of unobserved groups.

Using the censored normal likelihood function to define $P^j(y_it)$, the link function used to associate time with the outcome of interest is given by the latent variable y_{it}^{*j}:

$$y_{it}^{*j} = \beta_0^j + \beta_1^j Year_{it} + \beta_2^j Year_{it}^2 + \beta_3^j Year_{it}^3, ..., + + \beta_n^j Year_{it}^n + \varepsilon_{it} \tag{3}$$

where $Year_{it}$... $Year_{it}^n$ are the observed period, period squared, cubed and so on for each location, counting from 1 to 9 across the nine study periods we observe in our data, spanning 1940 and 2019; ε is the standard zero-mean and constant variance error term; and the various βs represent the parameters that determine the shape of each trajectory j. Given that these are group-specific, the model allows trajectories to differ markedly across groups.

For intuition on the approach, consider a hypothetical longitudinal distribution of regions, in which, unknown to the analyst, two groups exist, both tracing distinctive patterns of income growth over the study period. The first group consists of cities that have low average incomes in 1940, but then gradually become relatively rich by 2019. The second group starts out rich in 1940 but falls below the mean by 2019. A conventional assumption would be that the relationship between time and income levels is common to all locations, leading us to incorrectly conclude that relative incomes are static over the study period. Since GBTM treats the distribution as a mixture of unobserved groups, we would instead (correctly) detect two distinctive trajectories.

Once trajectories and probabilities of group memberships have been identified, and groups are described against a set of covariates, a further task is to explore theoretically-derived hypotheses about what makes a location likely to belong to a given group. This is undertaken using a multinomial logit model, in which, all else equal, initial variation in features of regional economies make one significantly more likely to be a member of group j versus a predetermined comparison group. For example, we can more formally test conjecture in Glaeser (2005) that regions that have a more highly educated workforce are more likely to follow resiliently high-income trajectories. The estimation of these 'risk factors' is undertaken at the same time as the group estimation, to incorporate the fundamentally probabilistic nature of group assignment.

We implement the group-based trajectory modeling approach as follows. The first step is model selection, in which the primary aim is to determine the number of latent groups best supported by the data. This is an iterative process, involving comparison of formal model fit statistics as well as expert judgment (Nagin, 2005). It also requires determination of the polynomial order that gives shape to each group's trajectory. Second, after settling on the optimal number of groups and polynomial order, we estimate posterior probabilities of group membership that support description of profiles and enable further model diagnostics. Third, we turn to hypothesis testing, in which we estimate the relationships between group membership and a series of location-specific predictors set to 1940 levels.

4 Data

The primary information to be used in this paper comes from a series of public-use microdata drawn from U.S. Census Bureau population surveys, compiled, harmonized and made available by IPUMS (Ruggles et al., 2021). While there exist other sources of information describing features of economic activity at the level of subnational regions in the United States, notably the Bureau of Economic Analysis' Regional Economic Accounts, none track the evolution of such indicators before 1969. To minimize bias, for each cross-section we exploit the largest available data extract. Effectively, this means we use the full count of the 1940 Decennial Census; one percent samples for 1950 and 1970; five percent samples for 1960, 1980, 1990, and 2000; a three-year, three percent sample of the American Community Survey (ACS) for 2010 (2009-2011); and a one-percent ACS sample for 2019. An alternative, 5-year, 5 percent ACS sample (2015-2019) is used as a robustness check on the 2019 data.

The 1940 Census was the first in which respondents were asked about their income, hence acts as a backstop before which we cannot directly estimate levels of interregional income inequality. We stop the modeling exercise at 2019, out of concern that the start of the COVID-19 pandemic may distort subsequent patterns. Throughout the study period, in-sample individuals are defined as those who do not reside in group-quarters; are not in schooling at the time of the survey; are between the ages of 16 and 65; and have nonzero income.

4.1 Geographic units

The primary spatial unit of observation in this study is the commuting zone (CZ). Commuting zones are groups of counties that are linked through the intensity of travel patterns and distinguished by weak inter-area commuting (Tolbert and Sizer, 1996). Commuting zones are not reported in Census data, hence we must assign individual respondents to them. We match individuals to commuting zones probabilistically, based on the smallest publicly identifiable geography in each Census. We assign each of these basic geographies a likelihood of belonging to each commuting zone, based on the population fraction in that commuting zone. Many locations map directly onto a single commuting zone. For individuals in locations for which multiple commuting zones are possible, we replace each observation with a multiple reflecting the number of potential commuting zone units to which each individual may belong. These receive adjusted person weights that reflect the likelihood that they reside in a given commuting zone. In other words, individuals are split into components whose size depends on the odds of living in a given commuting zone 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. Extending the procedure in Dorn (2009) to describe patterns in 1940, 1960, 2010 and 2019 results in 722 consistent, contiguous 1990-vintage Commuting Zone units that cover almost the entirety of the lower 48 states.[3]

4.2 Income Variables

In much of the analysis in this paper, the primary indicator of inter-place variation in incomes is the local mean of individuals' annual wage and salary income, which we deflate to constant 2015 dollars using the national urban consumer price index (CPI-U).[4] In some of the analysis, we detrend these income data, expressing them relative to the national mean. In keeping with the literature on spatial inequality (i.e., Manduca, 2019; Gaubert et al., 2021), this detrending permits us to highlight distributional changes that might otherwise be overwhelmed by broad-based, secular growth in incomes. We additionally describe estimates of household income, deflated to account for differences in local living costs. To accomplish this, following the approach described in Moretti (2013) and Kemeny and Osman (2018), we build a time-varying local consumer price index that combines three elements: local differences in the cost of housing, national non-shelter consumer costs, and national expenditure shares.[5]

One potentially complicating factor in measuring incomes and income inequality comes from the presence of topcoding in the underlying Census data. Our estimates indicate small but (somewhat unevenly) growing numbers of respondents are subject to topcoding on their wages in the public use data, ranging from a low in 1940 of 0.07 percent to a high in 2019 of 1.42 percent. This issue appears to be far more modest in relation to measures of total income. In the contemporary period at least, we expect topcoding to underestimate the degree of divergence for two reasons. First, there is likely more instances of topcoding in high-performing cities. Second, the quantity of income trimmed in topcoding is likely to be largest in these locations.

[3]Four commuting zone locations cannot be created in the 1940 data; these small locations are omitted from the analysis. We benefit from 1960 mappings made available by Evan K. Rose. See https://ekrose.github.io/resources/

[4]We also explore variation in total pre-tax income from all sources, which additionally includes income earned from business, welfare, social security, retirement and other sources. Household-level aggregations are also analyzed.

[5]If Y is unadjusted nominal household income for region j in time t, we obtain deflated real incomes R using the following formula: $R_{jt} = Y_{jt}/LCPI_{jt}$. The local consumer price index $LCPI$ is calculated as $LCPI_{jt} = (1 * w_{nt}) + [\frac{RENT_{jt}}{RENT_t} * w_h]$, where $w_n h$ and w_h represent expenditure shares derived from the CPI-U.

4.3 Covariates of local economic performance

We rely on various data sources to describe the distinguishing features of regions following common trajectories, informed by the literature surveyed earlier in this article. Variables are measured for commuting zones, and are organized into three broad categories of potential influence: urbanization; economic structure; and social structure.

To capture patterns of urbanization, we measure locations' population, aggregating county-level estimates from the County and City Data Book Consolidated File, made available as ICPSR 7736, as well as more recent information available directly from the Census Bureau. Since some studies measure urbanization using population density (Duranton and Puga, 2020), we additionally built indicators of population density using 2010 land area information from the TIGER/Geographic Identification Code Scheme.[6] Because population and density are so highly correlated (see Table A.1), we use only population in analyses. Population describes legacy effects of prior waves of development, capturing the inheritance from physical geographic features like navigable rivers, while also indicating scale effects that are seen to generate economic efficiencies in matching for workers, employers, buyers, and suppliers, as well as improved knowledge sharing.

Economic structure indicates regions' orientation towards certain forms of industrial activity. As we argued in our review above, across such a long study period, the ingredients for success can fundamentally change. To capture success in the earlier part of the study period, we measure the share of employment in manufacturing sectors, using workers' recorded industry of work in their main job, drawn from full counts and samples of the Decennial, as described above. In a similar way we also track employment shares in agricultural activities. Shifting to features that mark the more recent part of the study period, differences in local innovative effort are captured using geocoded and categorized data tracking the number of patents granted by the United States Patent and Trademark Office (USPTO), which we scale per 100,000 population. The data is drawn from HISTPAT (Petralia et al., 2016), which covers the study period up to 1975, thereafter using geographical patent information drawn directly from the USPTO.[7] Based on the presumed importance of the geographical clustering of highly educated workers (Autor, 2019; Card et al., 2021), we also use educational information in the Decennial and ACS to capture the share of workers who have attended at least four years of college.

We measure several factors shaping locational differences in social structure. To capture the institutions that enable and constrain educational attainment, using ACS and Decennial microdata we measure the share of workers that have less than a high school diploma. From these same data we estimate the proportion of the local labor force that self-identifies as black; and the share that is foreign-born. The latter indicates both the vibrancy of the local economy, but also its openness to diversity, considered to be linked to economic performance (Ottaviano and Peri, 2012; Kemeny and Cooke, 2018). Local shares of black workers capture a complex and key dimension of the American experience, involving historical patterns of settlement as well as internal migration and resettlement during the Great Migration of the 20th century (Boustan,

[6]See https://www.census.gov/quickfacts/fact/note/US/LND110210 for raw data.

[7]These data are available at the USPTO's PatentsView website: https://patentsview.org. Thanks to Sergio Petralia for sharing the cleaned data.

Table 1: Summary statistics on key variables

Variable	Mean	Std. Dev.
1940		
Annual Wages	15,842	3,861
Population (000s)	182	495
Population Density	56.1	171.7
Share Manufacturing	0.17	0.118
Share Agriculture	0.174	0.074
Share College	0.054	0.019
Patents per capita	3.768	8.399
Share Black	0.103	0.157
Share Dropout	0.679	0.083
Share Foreign-born	0.058	0.053
90/50 Wage Ratio	2.588	0.642
2019		
Annual Wages	49,540	7,430
Population (000s)	451	1,215
Population Density	117.8	301.4
Share Manufacturing	0.130	0.061
Share Agriculture	0.031	0.028
Share College	0.282	0.078
Patents per capita	20.78	36.03
Share Black	0.055	0.085
Share Dropout	0.053	0.025
Share Foreign-born	0.084	0.073
90/50 Wage Ratio	2.25	0.212

Note: Annual wages are deflated to constant 2015 dollars. Other measures as described in Section 4.3.

2016). Using worker-level information in these data, we capture inequality by measuring the ratio of incomes at the 90th to 50th percentiles of the local income distribution. Table 1 provides summary statistics describing 1940 and 2019 for local average incomes, as well as these covariates.[8]

5 Results

5.1 Overall trends in spatial income disparities, 1940–2019

We set the stage by providing a birds-eye view of the system. Over the period 1940-2019, Figure 1 describes the evolution of interregional income inequality across local labor markets in the United States. Panel A shows a series of (nearly) decadal population-weighted Gini coefficients, estimated on various measures of either individual or household income. For each indicator, Figure 1A reveals a long convergence period between 1940 and around 1980, followed by a reversal of this pattern. Trends in total and wage income track each other very closely.[9] The

[8] A correlation table for key variables is in Appendix A.

[9] Results are consistent when we exclude locations in the South, defined as South Atlantic, East South Central, and West South Central Census Regions. The remain consistent if we generate separate estimates for men and women, and if we restrict the sample to workers who are full-time, full-year employed. The shape of results is also

household series are also broadly comparable, deviating only in terms of a spike in inequality between 2010 and 2019. Adjusting for differences in local living costs do not fundamentally alter the aggregate patterns of spatial inequality.

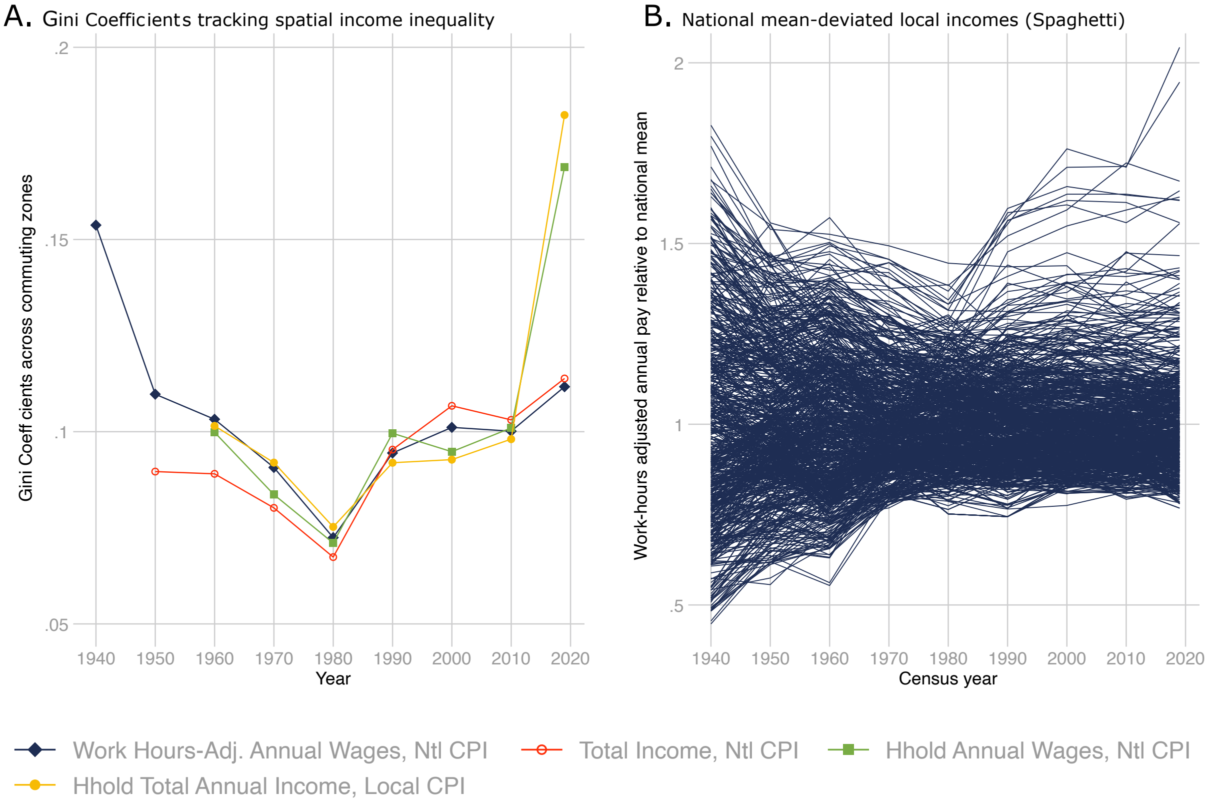


Figure 1: The fall and rise of spatial income inequality in the United States, 1940 to 2019
Note: Lines in Panel A are successive Gini coefficients, estimated on 722 commuting zones, built from public use Census decennial and ACS microdata, as described in the Data section of this paper. 'Hhold' stands for households. Income series are deflated either using the national urban consumer price index (CPI-U), or from a 'local' CPI, whose construction is described in Section 4.2. Each line in Panel B represents a single commuting zone. Incomes in Panel B are calculated as the ratio of local to all-location mean annual wage and salary income.

A more granular view emerges from the 'spaghetti' plot shown in Figure 1B. For each of the 722 commuting zones, this graph traces the evolution of local average hours-adjusted annual wages, detrended against the year-specific all-locations mean. The wasp-waisted pattern mirrors the decline and subsequent reversal of interregional wage gaps shown in Figure 1A. But it also provides a more textured picture of those trends. In statistical terms, most studies aim to describe the second moment of the distribution, while Figure 1B also yields insights into changes in its skewness, kurtosis and modality. It shows that, though the range of relative incomes is larger in 1940 than in 2019, the distribution is initially relatively uniform across that range. After 1940, convergence is evident, the product of a relative decline at the top, as well as the climbing up of initially poor locations. Then, starting in 1980, a bimodal pattern emerges, as a small subset of prosperous regions pulls up and away from an otherwise converging distribution. The recent divergence appears to be driven by a particular group of high-income regions; the

unchanged when we estimate β-convergence using 'Baumol regressions' (1986). All of these alternative estimates are available upon request.

Table 2: Highest income regions in 2019, with various measures of national economic significance

		% of U.S. Total		
Location	Relative Wages	Population	Employment	GDP
San Jose, CA	2.10	0.8	0.9	1.9
San Francisco, CA	2.00	1.6	1.8	2.9
Washington DC	1.72	1.2	1.6	1.9
New York, NY	1.69	3.7	4.2	5.9
Boston, MA	1.68	1.7	2.0	2.3 2
Newark, NJ	1.6	1.9	2.0	2.3 2
Tom's River, NJ	1.67	0.4	0.3	0.3
Bridgeport, CT	1.60	1.1	1.1	1.3
Seattle, WA	1.60	1.5	1.6	2.2
Baltimore, MD	1.51	0.8	0.9	1.0
Total	–	14.7	16.4	22

Note: 'Relative wages' means the ratio of local average wage and salary income to the average across all commuting zones. Each commuting zone is labelled according to its largest Census Designated Place (CDP). Hence, the commuting zone labelled 'San Francisco' is named for its largest CDP, but actually includes the following counties: Alameda, Contra Costa, Marin, Napa, SF, San Mateo and Solano. Whereas 'San Jose' covers Santa Clara, San Benito, Santa Cruz, and Monterey counties. Measures of population, employment, and GDP are authors' estimates built from county-level data drawn from the Regional Accounts of the Bureau of Economic Analysis.

range of income disparities was larger in 1940 than in 2019, but was much more evenly spread across the entire array of groups. Underneath the aggregate turn towards the 'great divergence', then, we observe a fundamental shift towards a polarization between a majority of locations that have experienced long-term convergence to the mean, and a small subset of commuting zones in which the average resident (though, as noted in Florida (2017), not every resident) has increasingly higher earnings as compared to those living elsewhere.

Are these conteporary high-performing cities merely outliers? This is a nontrivial question; if the answer is yes, then, rather than a narrative premised on rising spatial inequality and exclusion that demands remedy (as in, for instance: Austin et al., 2018), we should instead emphasize the convergence evident in the bulk of the urban system. To explore this idea, consider the ten commuting zones listed in Table 2, which in 2019 had average wages at least 1.5 times the all-location average. This threshold is arbitrary, although examining Figure 1B, one can discern that these recent high-performers constitute a relatively distinct group. For context, Table 2 includes data from the Bureau of Economic Analysis that describes each region's share of national employment, population and gross domestic product. The table makes clear that, while these are only 10 locations among 722, the temptation to consider them to be outliers should be tempered by their significance in the larger national economy. The group includes major metropolises, including ones centered on New York City (which includes Newark, Tom's River, and Bridgeport), the Bay Area, Seattle and Washington DC. In 2019, they contained over 47 million individuals, while their combined GDP made up fully 22 percent of the nation's. These are not therefore outliers; they represent a subset of places that have durably diverged from the rest of the system. Hence, the decision to weight Gini coefficients or other aggregate measures of inequality captures the relative importance of this subset of locations, but also qualifies the meaning of the contemporary inequality: in welfare terms, it means something different if the group enjoying very high relative incomes is or is not highly populated.

5.2 Turnover and persistence of regional income ranks

Does being a high-income region in time t predispose a place to remain at the top in time $t+n$? Note that the duration of the study period spans a world war; three major recessions; a five-fold growth in global trade as share of global economic output; and an epochal transition from the second (manufacturing-based) and third (information-based) industrial revolutions – each of these with geographically-differentiated effects. Amidst these shocks then, basic questions to answer include: is local prosperity in 1940 a predictor of prosperity in 2019? And to what extent is underperformance a trap?

Figure 2: Persistence and turbulence in (standardized) income ranks, 1940 and 2019
Note: N=722 Commuting Zones. Markers are scaled according to 2019 population. 95 percent confidence interval shown for solid linear fit line. Spearman's ρ = 0.48 (p=0.000). Dashed line drawn at 45 degrees.

Figure 2 displays the standardized rank correlation across these two periods. To aid interpretation, it also includes a dashed 45 degree line, as well as an unweighted linear fit line with a 95 percent confidence interval. Commuting zones are scaled according to their population in 2019. A positive correlation is evident, with a Spearman's ρ = 0.48. This suggests a degree of stability in the distribution: richer places in 1940 tend to be richer in 2019, and the typical poorer location tends to remain poor. Yet there is also substantial variability around the regression line. Many locations have substantially improved their position in the income hierarchy; and many others have slipped down. A set of populous regions appear at top rungs of the ladder in both periods, their persistence highlighted by their being centered directly over the 45 degree line.

Table 3 investigates this variation more systematically, via a transition matrix over quartiles

Table 3: Regional transition matrix across quartiles of average, hours-adjusted annual wages between 1940 and 2019, U.S. Commuting Zones

		2019 Income Quartiles				
		Q1	Q2	Q3	Q4	Total
1940	Q1	47%	31%	14%	8%	100%
Income		(85)	(56)	(26)	(14)	(181)
Quartiles	Q2	26%	30%	29%	14%	100%
		(47)	(54)	(53)	(26)	(180)
	Q3	16 %	28%	35%	22%	100%
		(28)	(50)	(63)	(40)	(181)
	Q4	12%	12%	21%	56%	100%
		(22)	(21)	(37)	(100)	(180)
	Total	25%	25%	25%	25%	100%
		(182)	(181)	(179)	(180)	(722)

Note: Q1 represents lowest income quartile, Q4 is highest. Percentages are rounded to the nearest whole number. Actual counts in parentheses.

of average, hours-adjusted annual wages. The table captures the likelihood that a region occupying a certain quartile of the distribution in 1940, will have, by 2019, meaningfully changed its position. The largest shares in Table 3 are all along the diagonal, representing locations that remain fixed in the same quartile in both 1940 and 2019. And yet, in the most stable quartile, only slightly more than half the regions that were in the highest quartile (Q4) in 1940 remained there in 2019. In short, despite there being a degree of persistence across these broad groupings, many regions appear to have substantially changed their fortunes over this period. Perhaps unsurprisingly, it is those regions in the middle two quartiles are those most likely to have done so. Those in the lower-middle quartile (Q2) in 1940 were 17 percent more likely to move into one of the two upper quartiles than to move down, indicative of the powerful convergence process visualized in the right panel of Figure 1. Meanwhile, by 2019, almost half the locations that sat in 1940 in the upper middle quartile (Q3) fell into the lower half of the distribution. The wider conclusion to be drawn from this table is that the system displays patterns consistent with both persistence and turbulence; the former most sharply at the tails.

5.3 Development trajectories of regions: Group based trajectory modeling

Using the GBTM tecnniques described in Section 3, we aim now to unpack aggregate patterns described thus far, by identifying a set of distinctive, long-run pathways of development traced out by regional economies. One way to think about the GBTM analysis is that we seek to boil down the many strands of spaghetti visualized in Figure 1B into distinct 'clumps', with each clump representing a group of local labor markets. Group membership is defined in terms of the coherence over the 1940-2019 period of trajectories of local average incomes relative to the all-locations mean.

Having explored solutions that assign the 722 locations to between one and ten groups, formal goodness-of-fit measures including the Bayesian Information Criterion (BIC), as well as

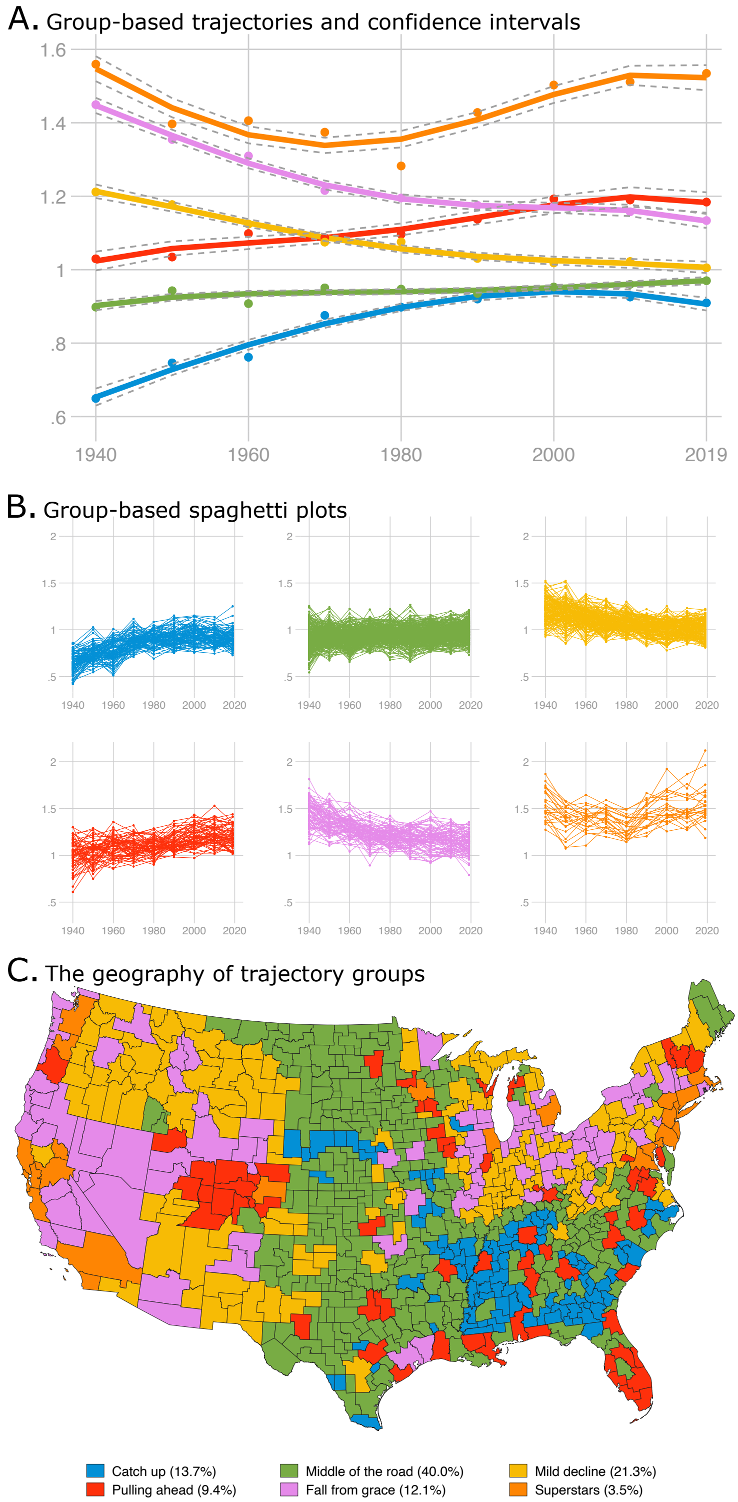


Figure 3: Trajectory groups in the evolution of wages 1940–2019
Note:N=722 Commuting Zones (CZ). Trajectories estimated using a group-based trajectory model with the censored normal distribution and a polynomial order of {3,4,4,4,4,4} across groups numbered 1-6. Dotted lines in Panel A represent 95 percent confidence intervals. Y-axes in Panels A and B measure the ratio of local to national average annual wages. Percentages indicate the odds that a randomly-selected CZ will follow a particular trajectory.

substantive expertise, favor a six-group model.[10] Figure 3A displays the resulting trajectories, along with confidence intervals, with the legend (at the bottom of Figure 3) representing the probability, π, that a randomly selected commuting zone will follow trajectory group j. The figure indicates that trajectory groups follow distinctive paths over the study period. Confidence intervals are uniformly tight, suggesting precision in the estimated trajectories. Post-estimation diagnostics and various robustness checks indicate a stable model in which group assignment is obtained with a high level of accuracy. For instance, the trajectories' average maximum posterior probability ($AvePP_j$) ranges from a low of 0.929 to a high of 0.997. For trajectory j, an $AvePP_j$ of 0.929 indicates that the average region in this trajectory has a 93 percent likelihood of being correctly assigned to that trajectory. Values obtained for $AvePP_j$ are uniformly close to one, which suggest very high levels of certainty regarding classification. Further, trajectories remain strikingly similar even when we limit the sample of regions to those with initial (1940) populations of at least 100,000.[11]

Regions likely to follow what we label as the 'Catch-Up' trajectory were considerably poorer than the national average in 1940, but climbed to approximately 90 percent of the mean by 1980. The 'Pulling Ahead' group captures a less common pathway, starting near the national average in 1940 and then progressing by 2019 to approximately 1.2 times the mean. A randomly-drawn region has a high probability (40%) of following what we call 'Middle of the Road' path, which hovers slightly below the mean throughout the study period, very gradually approaching it. The 'Mild Decline' group is the inverse of the 'Pulling Ahead' group: locations in the former group experience a moderate downward pattern in terms of average wage income relative to the national mean. Meanwhile, some regions have experienced a 'Fall from Grace': in 1940 they were highly prosperous, with initial average incomes more than 1.4 times the national mean, but over the study period they tumble downwards towards the mean. The final group are the 'Superstars': an atypical path that consists of places that have resiliently been high performers over the entire duration of the study.

Figure 3B visualizes variation within the groups, bridging the inter-group group trajectories and the 722-zone spaghetti plot (Fig.1B). Figure 3B confirms the high degree of coherence of these groups. Each sub-panel tracks a set of individual trajectories that are both relatively internally homogeneous and externally differentiated. While individual locations may stray from dominant group patterns, these deviations remain brief. Based on maximum posterior probabilities, Figure 3C maps the members of these different trajectory groups in the continental United States.

5.3.1 Describing trajectory groups

To begin exploring group differences, Figure 4 describes the evolution of each group identified in the GBTM analysis in terms of indicators of urbanization, economic and social structure. To facilitate comparison, values in the figure are presented as deviations from all-location means. Initial levels of the three groups of variables are likely to have influenced subsequent development, but national trends have also shifted, representing structural changes that shape the

[10]See Appendix B for model fit statistics and other details on model selection.

[11]This reduces the number of regions from 722 to 297. Consult Appendix B for a report on this and a battery of other diagnostics and robustness checks.

fates of all regions. For reference, over the study period, all-location means for the following characteristics of economies increased: population, college education, innovation, immigration and income inequality.Meanwhile, all-location means declined in manufacturing and agricultural employment shares, high school dropout rates, and the proportion of Blacks in the workforce.[12]

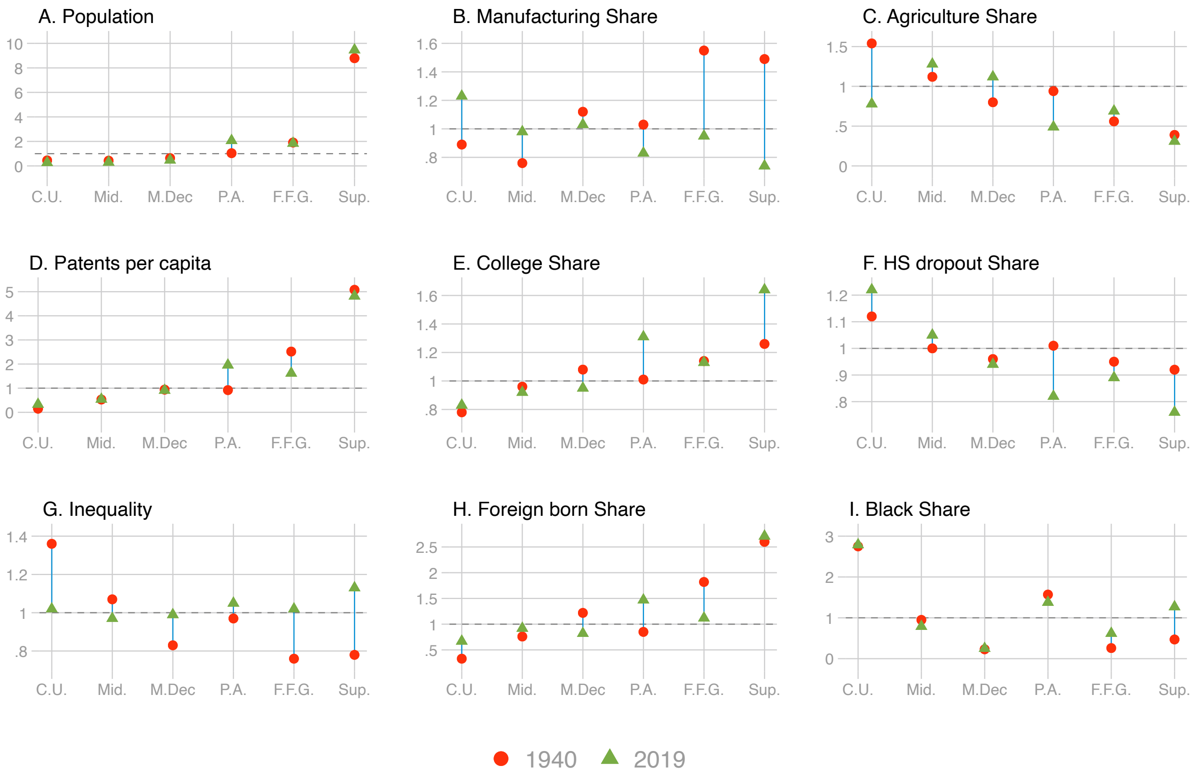


Figure 4: Describing trajectory groups using mean deviations
Note: 'C.U.' ='Catch-Up' group; 'Mid' = 'Middle of the road'; 'M.Dec'= 'Mild Decline'; 'P.A.'='Pulling ahead'; 'F.F.G.'='Fall from grace'; and 'Sup.'='Superstars'. Statistics presented are deviations of individual group means against the overall mean for all commuting zones (equal to 1 and designated using a horizontal dashed line). Inequality is the ratio of annual wages at the 90th percentile against the 50th percentile. Patent counts in the second period are for 2010 – the closest period in which sufficiently geographically-detailed data are available. For all variables except patenting, the underlying data come from the full count 1940 Decennial Census and a one percent extract of the 2019 ACS. Patent data from HISTPAT. Data details are in Section 4

Considering Figures 3C and 4, Catch-up regions consist primarily of formerly low-income Southern locations that benefited from the postwar spatial economic integration of the U.S. This integration was partly achieved by the movement of industry south from the late 1940s onward, allowing these regions to reduce their dependence on agriculture even faster than the nation as a whole and, against the national trend, to shift strongly toward manufacturing (Rees and Norton, 1979). This also generated migration flows from rural to urban areas, as well as from north to south. Development economics predicts that catch-up is typically more rapid from a low-income starting point, which avers true in this case. But, in spite of initial progress, in the longer-run, their small populations, limited expansion in educational endowments and modest patenting could be taken as limits to the further takeoff of this group.

[12]Appendix Figure B.3 depicts these 'national' trends

Middle of the Road economies are mostly found in the Midwest and South. They have remained close to U.S. mean income over many decades of significant structural change, suffering neither great declines nor enjoying soaring incomes and population growth. They have maintained shares of agriculture and manufacturing. Meanwhile, like the Catch-Up group, their growth in the high-income fundamentals of the 21st century economy – patenting and graduate density – has been below average.

Mild Decline locations are relatively small in population, generally found in the Intermountain West and industrial Northeast, as well as Arizona and New Mexico. Most of their fundamentals resemble those of Middle of the Road regions, but they have divested more strongly from manufacturing. However, growth in college educated workers for this group has increased more slowly than in the national population, while growth in patents tracks the national mean. These locations have also fallen behind in their ability to attract immigrants; and they are less Black than the national average.

Pulling Ahead regions are moving up the income hierarchy, while experiencing faster-than-average population growth. Their geography is striking. Unlike Superstars, they are not bi-coastal, and are mostly found in the metropolitan South, the Colorado plateau, and Florida, but there are a few in the upper midwest as well. The contrast with Mild Decline locations is stark. Pulling ahead regions such as Atlanta, Dallas and Austin show striking increases in population, patenting, and shares of college graduates and immigrants, all at rates that considerably outstrip national growth. Meanwhile, they have dramatically reduced dropout rates and include relatively high-share Black populations.

Fall from Grace regions include many former leading lights of the American economy during the manufacturing era, including Cleveland and Pittsburgh, as well as some resource-intensive Western regions. The former are large city-regions whose populations grew considerably during the industrial-urban transition of the late 19th-early 20th centuries, fuelled then by both domestic and international migration. The sharp decline in their manufacturing base has not been replaced by the features that strongly mark the Pulling Ahead places: strong attraction of the new generation of immigrants, increasing shares of college graduates and rising patenting. In other words, though these regions start out with some evident strengths, they do not renew such advantages over time, suggesting that their initial strengths in education, immigrant attraction and innovation could not be translated from the knowledge-based industries of the past to those of the present. This result suggests, in contrast to some findings in the literature (i.e., Glaeser, 2005), that education is not enough to assure prosperity over the long run.

Superstar regions follow a trajectory that contrasts sharply with Fall from Grace locations, but with several similar initial characteristics. Superstars also experienced dramatic declines in manufacturing, but they transitioned successfully to the new economy. Unlike the Fall from Grace regions, Superstars improved their already high college graduate shares at a rate that outstripped the five-fold national increase. Meanwhile, their already high initial levels of patenting and immigrant attraction kept pace with major national increases. All of this came with greater-than-average expansion in population. As a consequence, though the Superstar group consists of only 25 out of 722 commuting zones, data from the Bureau of Economic Analysis' Regional Economic Accounts indicates that, in 2019, their footprint on the country was the largest of all the groups, hosting 32 percent of the national population and 41 percent of its Gross Domestic Product. Regions such as New York, Los Angeles and Boston have been large, successful city-regions for a long time, but the degree to which they have outperformed other locations has changed over the period under scrutiny, with a nadir during the mid-century leveling of interpersonal and geographical inequality, then rising again from 1980.

Several wider narratives emerge when we synthesize these individual associations. First, over the study period, having a large population appears to be increasingly linked to sustained prosperity, potentially as a cause and certainly as an outcome. This represents a challenge for smaller places in the future, although – as we will discuss – large populations are not sufficient to ensure prosperity. Futher, developing an economic structure around the key skilled and knowledge-intensive activities of the day descriptively differentiates Pulling Ahead and Superstar regions from those following a Fall from Grace trajectory. Increasing manufacturing helps Catch-up places to rise to middle-income ranks, in a way that is directly analogous to pathways followed by lower-income countries in relation to the world economy. Educational attainment – measured at both ends of the distribution – moves in a manner consistent with the common notion that, in an economy increasingly based on knowledge, localities will be best positioned when they have greater endowments of well-educated workers. Stocks of human capital may not prevent occasional downturns, but, as in Glaeser (2005), it may ensure that they are transitory periods within a broader pattern of long-term successful resilience. Meanwhile strengthening social structures, and especially low shares of high school dropouts and openness to immigration are associated with relative success.

The use of incomes as the defining characteristic of group membership results in some within-group heterogeneity – a result that would hold true for individual or combinations of characteristics. Our focus on income is justified by its robustness as a proxy for the quality of economic development, but growth also contains other, at least partly independent dimensions, notably quantity, which we might approximate using population growth.[13] To take an example, the Pulling Ahead group consists of familiar cases in the South that combine strong growth in both population and incomes, including Atlanta, Orlando, Austin, Charlotte, Dallas, Fort Worth and Nashville. But this group also includes New Orleans, which in the popular imagination is not commonly grouped with the others. Schematically, for some regions, there appear to be strong trade-offs between quantity and quality, whereas for others the two are more closely aligned. New Orleans is a case of the former, whereby average regional incomes have grown in the context of significant population decline.

We observe some similar patterns in other groups. Among the Superstars, for instance, alongside the expected coastal global metropolises, we find Baltimore, Detroit and Philadelphia. These are city-regions with long-term success in maintaining average incomes, against a background of sluggish population growth. The Fall from Grace group offers a particularly rich set of surprises that merit future analysis. These include cases where starting points in 1940

[13]Exploring this, Figure C.1, in Appendix C, shows trends in population shares for each trajectory, going all the way back to 1900. It is striking that the two groups with high or improving income ranks (Superstars and Pulling Ahead) are the places whose shares of national population have grown considerably over the study period, in contrast to all the other groups. Income, at first glance, seems a powerful attraction for people, although causality may be due to the attributes of those who have sorted into these places.

consisted of small but relatively wealthy populations that by 2019 became utterly transformed, becoming large, middle-income cities. Cities such as Phoenix, Tucson, Las Vegas, Salt Lake City, Houston, Santa Barbara and Santa Fe appear to fit this pattern. Given their different initial income levels, cities in the Pulling Ahead and Fall from Grace groups might be examples of mean reversion in a process of urbanization. But this idea is weakened by our evidence. High levels of population growth do not automatically cause an urban area to decline in relative income, as in the examples above; in the post-1940 period, for example, Superstars such as Los Angeles and San Francisco added millions of people while rising up the income ranks of US metro areas, and Houston seems to be repeating this history. Indeed, some of the growing Fall from Grace areas might be hitting a long-term middle income trap. This possibility echoes a finding in international economics: that rapid initial income growth is commonly succeeded by stasis and stagnation (Eichengreen et al., 2013) – a phenomenon also described at a regional scale in Europe (Diemer et al., 2022). While some regions are mired in a middle income trap, some of the regions rising up income ranks may become future superstars, with Austin a seemingly good candidate. The varied interrelationships between the quantity and quality of growth deserve deeper study, notably to see what separates the cases of income preservation in the face of sluggish demographics; income growth with high population growth; income stagnation with high population growth; and income stagnation with sluggish growth or declining population shares. We do not yet have a systematic picture of these finer differences within and among groups; such a picture would add much to simple tradeoffs between population and income found standard spatial equilibrium frameworks, and are likely to disprove any simple relationship between population change and mean reversion in incomes.

5.3.2 Initial conditions and subsequent trajectories

The previous section described how each trajectory group has evolved over the study period, in terms of urbanization, economic structure, and social structure. We can think of trajectories as the result of an interaction between initial endowments of such characteristics, local and economy-wide shocks, and subsequent changes in these features. Existing theory provides little clarity on the nature of these interactions and their relationship to income. In the absence of an accepted model that would specify these relationships, and in the presence of endogeneity, we cannot offer a causal analysis here.

However, starting more modestly, what we can do is investigate the roles played by initial endowments of urbanization, economic structure and social structure in selecting regions into subsequent pathways of development. Over our study period, initial features represent snapshots of locations at a particular moment in time (1940). They also point to their differing deep roots – the outcomes of longstanding historical processes. That history may be important is a notion developed in both institutionalist (i.e., North, 1987) and evolutionary economics (i.e., Nelson and Winter, 1982). In these theories, certain starting conditions can hold a local economy back from developing or attracting emergent high-wage activities, while others favor reinvention and success.

To explore these themes empirically, we turn to multinomial logistic regression, which we estimate within the GBTM procedure as a means of incorporating the probabilistic assignment of regions to trajectories. The general form of question asked here is: controlling for other potentially-relevant selection factors, do initial differences in local feature z predict a region's probability of belonging to trajectory x versus a 'reference' trajectory y? Across the structural shocks of the economy as a whole and changing patterns of convergence and divergence, this enables us to explore how initial conditions or deep roots of an economy predict whether it will follow a particular pathway.

Table 4: Predictors of place membership in groups of income trajectories (Middle of the Road is reference group)

	Catch Up	Mild Decline	Pulling Ahead	Fall from Grace	Superstar
Population	0.14***	2.69**	3.55***	17.83***	20.00***
Manufacturing	1.08	0.97	0.92**	0.92*	0.86**
Agriculture	1.42***	0.78***	0.93	0.77***	0.71**
Patenting	0.67*	1.03	1.02	1.08	1.1*
Graduates	0.95	1.35*	1.00	1.55*	1.07
Dropouts	1.05	1.06	0.98	0.87	0.72**
Inequality	17.97***	0.00***†	0.02***	0.00***‡	0.00***††
Foreign-born	0.74***	1.28***	1.24***	1.48***	1.78***
Black	1.01	0.89*	1.12***	1.10	1.29**

Note: Odds ratios with significance stars shown, where * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All predictors in the model are set to absolute values in 1940. Population is measured in 100,000s; Manufacturing, Agriculture, Graduates, Dropouts, Foreign-born and Black are shares of total employment (scaled 0-100); Patenting is measured as the number of granted USPTO patents per 100,000 population; Inequality is the ratio of the local 90th percentile of income to 50th. Actual odds ratio for †: 0.000706476937277; actual odds ratio for ‡: 0.000000000167976; actual odds ratio for ††: 0.000000001533106. Full regression results available in the Appendix, Table B.2.

Table 4 presents odds ratios and significance levels from this model.[14] Throughout, we use the Middle of the Road trajectory as the reference group, on the basis that its average income levels most closely track national averages throughout the nearly 80-year study period. This is also the group with the largest likely membership (in terms of number of territories, though not overall share of the population). This group thus represents both a 'common' location type, as well as one less marked by sharp convergence or divergence, against which to benchmark deviations.

Odds ratios above one in this analysis imply that a one-unit increase in predictor z raises the likelihood that a location follows a particular trajectory instead of the Middle of the Road pathway. Odds ratios below one indicate that higher values of the predictor make it more likely that region will instead follow the Middle of the Road trajectory. The odds ratios can be compared across groups. To illustrate, a location with an additional 100,000 people in 1940 will be more likely to be in the Superstar group as compared to the Middle of the Road group, as will an additional percent in initial foreign-born population. However, the marginal 'effect' of this increase in population on the odds of following the Superstar path is larger than for an additional percent share of immigrants in the local workforce.[15]

[14] Coefficients and standard errors are presented in Appendix B

[15] While in a linear regression, generating standardized z-scores might be a sensible approach to maximizing comparability, such normalization techniques reduce predictive accuracy in logistic regression (Turing, nd), while

Assuming a threshold for statistical significance of 5 percent, Table 4 suggests that regions 'select' out of the Middle of the Road pathway and into the Catch-up trajectory by having smaller initial populations, being more focused on agriculture, less innovative, more unequal, while having smaller proportions of foreign-born workers. One framework for understanding for this contrast, coming from the field of international development, is the classical idea of the "advantages of backwardness" (Gerschenkron, 1952), in which, in response to growing integration and investment, low initial levels of income facilitates high rates of growth, pushing them toward the mean. Some of these are obviously rural areas that began to urbanize, such that declining agricultural employment, on the margin, was positively associated with income growth. What an initial specialization in agriculture does *not* do is assure further growth beyond a certain point, as attested to by abundant evidence of a middle income trap (Eichengreen et al., 2013); and given that most of the US population now lives in clearly non-agricultural areas, it will not explain much about the other income dynamics we have detected.

In the contrast between Mild Decline and Middle of the road, we find that initial population, shares of college graduates and foreign-born workers select positively into the former group, while agriculture and Black share select negatively. Pulling Ahead regions are initially: larger, more African American, and more immigrant-dense, while being more egalitarian, and less focused on manufacturing. Interestingly, initial patenting activity does not significantly differentiate Pulling Ahead regions from Middle of the Road locations – an outsize propensity to innovate was incorporated into the former set of locations at some point after 1940, but was not initially a distinguishing feature.

The probability of following a Fall from Grace pathway versus Middle of the Road increases with higher initial levels of population, college graduates and immigrants. It declines by having an initially stronger focus on manufacturing and agriculture. Finally, significant predictors of following a Superstar pathway are a large initial population, higher proportions of immigrant and Black workers, lower shares of manufacturing and agricultural employment, fewer dropouts and higher levels of patents per capita. Though inequality emerges as statistically significant, the diminutive size of the odds ratios suggests a negligible role in selection across all but the Catch Up group. One possible interpretation of this finding is that, for these other groups, interpersonal income inequality is a byproduct of development, rather than a structural determinant.

6 Conclusion: The many faces of divergence and convergence

This article demonstrates that, since 1980, the American urban-regional system has been marked simultaneously by income convergence and divergence. This finding contrasts with most of the literature, which emphasizes only either aggregate divergence or the end of convergence. Income levels in a small number of superstar cities are indeed increasingly pulling away from the rest of the national system, but the rest of that system continues a secular process of upward convergence evident since at least 1940.

We also distinguish this specific, polarized form of today's American spatial inequality from

also reducing the comprehensibility of odds ratios.

the income disparities of the past. In 1940, we show that much of the national variation in income levels was due to the existence of low-income regions in a much less integrated national economy, notably in the South, Appalachia and the interior West. Incomes then were also more evenly spread across the system. Today's bifurcated divergence occurs in a much more integrated national economy, where some of the historical pockets of low income, notably those in the Deep South, have been eliminated, while at the same time, new pockets of possible stagnation have emerged in formerly prosperous regions.

Nonetheless, Superstars are not merely outliers that might be discounted against an urban system dominated by the forces of convergence; in 2019, these cities account for 32 percent of the national population and 41 percent of its Gross Domestic Product. They have been high-performers across study period, but the magnitude of their divergence and their weight in the national economy has changed. This trajectory is an example of positive persistence: these are metropolitan regions that have successfully reinvented themselves over the study period, capturing income-enhancing aspects of structural changes in the wider economy, while shedding those elements that act as a brake on growth.

For all other trajectory groups, there is considerable turbulence or churn in positions. Some regions have emerged from relative poverty, capturing industrial activities from previous industrial revolutions that have spatially decentralized, and thereby yielding income growth that allows them to catch up to the rising national mean. For some other regions, across the dividing line of 1980, past success has not translated into continued strong economic performance.

Consider, for example, the growing importance of innovation in shaping American prosperity. In 1940, the average superstar location generates around six times as many patents as the all-locations average, thereafter keeping pace with a nearly six-fold national increase. Among regions that are initially low in patenting activity, some manage to increase their rate of innovation faster than the trend, and others are unable to keep pace; still others – specifically regions in the Fall from Grace group – fail to maintain their initial strength in innovation. Applying a logic from Boschma (2015) and other evolutionary economic geographers, it may be that Superstars and the Fall from Grace groups differ in terms of opportunities to build on previously high levels of innovation, adapting them towards new (and related), promising technological fields. Abstracting, for some regions and indicators, initial positions are a source of continuity; for others, initial values appear unrelated to subsequent performance; beyond initial positions then, there are additional selection forces at work.

Further understanding these dynamics is a major area of future research opened up by the present investigation. As we have done in this article, there is promise in pursuing research that integrates systemic and place-centered perspectives. The tools applied here can be extended to gain deeper insight into both the system as a whole and the developmental trajectories of places. One valuable exercise would be to use groups we have identified as the basis for closer comparative analysis. For instance, we should urgently seek to answer why, from seemingly similar starting points in 1940, the Superstar and the big older manufacturing regions in the Fall from Grace groups have performed so differently. Such forensic analysis should generate greater clarity on the forces that drove two formerly successful sets of regions down markedly different long-run paths.

Whether or not there is aggregate convergence, the viewpoint from a region that is declining toward the mean looks entirely different from that of one ascending the ranks. Even if we compare regions that have similar income levels at a given moment in time, a downwardly-mobile region will be buffeted by forces not found in a city on the rise. As we learn from the European case, stagnant locations are likely to be suffering collateral damage from population outflows, declining social mobility, pressure on public budgets, shrinking property values and household wealth, and deteriorating social-economic networks (Diemer et al., 2022).

In closing, it should be noted that models of convergence and spatial equilibrium remain powerful in part because of the appeal of their assumptions regarding efficient and welfare-maximizing distributions of people and activity, in each case predicting that a specific kind of arbitraging behavior will produce a tendency towards equalization – either of productivity or welfare. The normative appeal of such outcomes help make these models powerful reference points for policymakers.

But are they the right frames for understanding today's spatial inequality? The last durable period of convergence, ending in 1980, exhibited features that may no longer mark the contemporary period. The 1940 to 1980 period witnessed considerable spread, both westward and southward, of people and economic activities. Multidirectional internal migration proceeded at a high level across all skill groups, complemented by the Great Migration of African-Americans. Meanwhile, major firms, their supply chains, and markets became increasingly nationally integrated. Except for the latter, these mechanisms that enabled convergence may no longer be as powerful. The economic geography of the United States in the 21st century is characterized by lower rates of long-distance migration; increasingly inelastic housing supplies in high income areas; increased sorting of people by skill across space; and enduring geographical pockets of under- and non-employment (Austin et al., 2018). Paradoxically, even though we observe that five out of six trajectory groups converge since 1980, the growing income divergence and increased demographic weight of Superstars is both durable, and has accelerated since 1980. Although mobility patterns during the COVID-19 pandemic suggested a major dilution of superstar drawing power, recent work highlights a strong rebound of domestic and international migration into superstar locations like New York City and San Francisco (Frey, 2023). Some of the pathways that we identify are shaped by the locational sorting and arbitraging that is the core of convergence thinking, notably in the rapid population growth of some areas. But there is little indication that this shift is sufficient to generate a shift toward either nominal or real income convergence. In light of the temporal durability and recent growth of the Superstars in relation to the rest, and the existence of some evidently problematic income trajectories, there is an ongoing need for further research on how these trends are related to market forces, institutional forces, and policies. Whether what we observe is a distortion away from convergence, or instead the normal workings of a modern space economy, has enormous implications for the kinds of policies we should adopt in attempting to shape the geography of the economy.

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, 1:385–472.

Austin, B. A., Glaeser, E. L., and Summers, L. H. (2018). Jobs for the heartland: Place-based policies in 21st century america. National Bureau of Economic Research Working Paper 24548.

Autor, D. (2019). Work of the past, work of the future. National Bureau of Economic Research Working Paper No.25588.

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. and Sala-i-Martin, X. (1991). Convergence across states and regions. *Brookings papers on economic activity*, pages 107–182.

Barro, R. J. and Sala-i-Martin, X. (1992). Convergence. *Journal of political Economy*, 100(2):223–251.

Baum-Snow, N. and Pavan, R. (2012). Understanding the city size wage gap. *The Review of economic studies*, 79(1):88–127.

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

Baumol, W. J. (1986). Productivity growth, convergence, and welfare: what the long-run data show. *The american economic review*, pages 1072–1085.

Boschma, R. (2015). Towards an evolutionary perspective on regional resilience. *Regional studies*, 49(5):733–751.

Boustan, L. P. (2016). *Competition in the Promised Land*. Princeton University Press.

Card, D., Rothstein, J., and Yi, M. (2021). Location, location, location. US Census Bureau Center for Economic Studies Discussion Paper 21-32.

Carlino, G. A. (1992). Are regional per capita earnings diverging? *Business Review*, 3:3–12.

Case, A. and Deaton, A. (2020). Deaths of despair and the future of capitalism. In *Deaths of Despair and the Future of Capitalism*. Princeton University Press.

Chapple, K. and Lester, T. W. (2010). The resilient regional labour market? the us case. *Cambridge journal of regions, economy and society*, 3(1):85–104.

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. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chinitz, B. (1961). Contrasts in agglomeration: New york and pittsburgh. *The American Economic Review*, 51(2):279–289.

Christopherson, S., Michie, J., and Tyler, P. (2010). Regional resilience: theoretical and empirical perspectives. *Cambridge journal of regions, economy and society*, 3(1):3–10.

Ciccone, A. and Hall, R. (1996). Productivity and the density of economic activity. *The American Economic Review*, 86(1):54–70.

Clark, J. and Bailey, D. (2018). Labour, work and regional resilience. *Regional Studies*, 52(6):741–744.

Clogg, C. C., Petkova, E., and Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, 100(5):1261–1293.

Colen, C. G., Ramey, D. M., Cooksey, E. C., and Williams, D. R. (2018). Racial disparities in health among nonpoor African Americans and Hispanics: The role of acute and chronic discrimination. *Social Science & Medicine*, 199:167–180.

Connor, D., Berg, A., Kemeny, T., and Kedron, P. (2023). Who gets left behind by left behind places? *Cambridge Journal of Regions, Economy and Society*, pages 1–27.

Connor, D. S. and Storper, M. (2020). The changing geography of social mobility in the united states. *Proceedings of the National Academy of Sciences*, 117(48):30309–30317.

Connor, D. S., Uhl, J. H., Xie, S., Talbot, C., Hester, C., Jaworski, T., Gutmann, M. P., Leyk, S., and Hunter, L. M. (2022). Rising community poverty reduces social mobility for rural children. Available at SSRN https://dx.doi.org/10.2139/ssrn.4127500.

Couture, V., Gaubert, C., Handbury, J., and Hurst, E. (2019). Income growth and the distributional effects of urban spatial sorting. National Bureau of Economic Research Working Paper 26142.

Cramer, K. J. (2016). *The politics of resentment: Rural consciousness in Wisconsin and the rise of Scott Walker*. University of Chicago Press.

Dauth, W., Findeisen, S., Moretti, E., and Suedekum, J. (2022). Matching in cities. *Journal of the European Economic Association*, 20(4):1478–1521.

Davies, S. (2011). Regional resilience in the 2008–2010 downturn: comparative evidence from european countries. *Cambridge Journal of Regions, Economy and Society*, 4(3):369–382.

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

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

Diamond, R. and Moretti, E. (2021). Where is standard of living the highest? local prices and the geography of consumption. National Bureau of Economic Research Working Paper 29533.

Diemer, A., Iammarino, S., Rodríguez-Pose, A., and Storper, M. (2022). The regional development trap in europe. *Economic Geography*, 98(5):487–509.

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. and Puga, D. (2020). The economics of urban density. *Journal of Economic Perspectives*, 34(3):3–26.

Durlauf, S. N. and Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of applied econometrics*, 10(4):365–384.

Eckert, F., Hejlesen, M., and Walsh, C. (2022). The return to big-city experience: Evidence from refugees in denmark. *Journal of Urban Economics*, page 103454.

Eggleston, E. P., Laub, J. H., and Sampson, R. J. (2004). Methodological sensitivities to latent class analysis of long-term criminal trajectories. *Journal of Quantitative Criminology*, 20(1):1–26.

Eichengreen, B., Park, D., and Shin, K. (2013). Growth slowdowns redux: New evidence on the middle-income trap. National Bureau of Economic Research Working Paper 18673.

Florida, R. (2017). *The new urban crisis: How our cities are increasing inequality, deepening segregation, and failing the middle class-and what we can do about it*. Hachette UK.

Frey, W. H. (2023). Pandemic-driven population declines in large urban areas are slowing or reversing, latest census data shows. Brookings, Retrieved May 28, 2023 from https://www.brookings.edu/research/pandemic-driven-population-declines-in-large-urban-areas-are-slowing-or-reversing-latest-census-data-shows/.

Galbraith, J. K. and Hale, J. T. (2014). The evolution of economic inequality in the United States, 1969-2012: Evidence from data on inter-industrial earnings and inter-regional incomes. *World Economic Review*, 3(2014):1–19.

Galor, O. (1996). Convergence? inferences from theoretical models. *The economic journal*, 106(437):1056–1069.

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

Gaubert, C., Kline, P. M., Vergara, D., and Yagan, D. (2021). Trends in US spatial inequality: Concentrating affluence and a democratization of poverty. National Bureau of Economic Research Working Paper No.28385.

Gerschenkron, A. (1952). Economic backwardness in historical perspective. In Hoselitz, B., editor, *The progress of underdeveloped countriesd*. Chicago University Press.

Glaeser, E. L. (2005). Reinventing Boston: 1630–2003. *Journal of Economic Geography*, 5(2):119–153.

Glaeser, E. L. (2008). *Cities, agglomeration, and spatial equilibrium.* OUP Oxford.

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

Gong, H., Hassink, R., Tan, J., and Huang, D. (2020). Regional resilience in times of a pandemic crisis: The case of covid-19 in china. *Tijdschrift voor economische en sociale geografie*, 111(3):497–512.

Grandin, G. (2019). *The end of the myth: From the frontier to the border wall in the mind of America.* Metropolitan Books.

Greene, W. H. (1990). *Econometric analysis.* Macmillan.

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

Houlden, V., Robinson, C., Franklin, R., Rowe, F., and Pike, A. (2022). Locating "left behind" places and people in england: Scale, trajectory, and the challenge of multidimensionality. OSF Preprints.

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

Jones, B. L. and Nagin, D. S. (2013). A note on a Stata plugin for estimating group-based trajectory models. *Sociological Methods & Research*, 42(4):608–613.

Jones, C. I. and Klenow, P. J. (2016). Beyond gdp? welfare across countries and time. *American Economic Review*, 106(9):2426–57.

Kaplan, G. and Schulhofer-Wohl, S. (2017). Understanding the long-run decline in interstate migration. *International Economic Review*, 58(1):57–94.

Kemeny, T. and Cooke, A. (2018). Spillovers from immigrant diversity in cities. *Journal of Economic Geography*, 18(1):213–245.

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.

Kemeny, T. and Storper, M. (2015). Is specialization good for regional economic development? *Regional Studies*, 49(6):1003–1018.

Kemeny, T. and Storper, M. (2020). The fall and rise of interregional inequality: Explaining shifts from convergence to divergence. *Scienze Regionali*, 19(2):175–198.

Kim, S. (1998). Economic integration and convergence: Us regions, 1840–1987. *The Journal of Economic History*, 58(3):659–683.

Klijn, S. L., Weijenberg, M. P., Lemmens, P., van den Brandt, P. A., and Lima Passos, V. (2017). Introducing the fit-criteria assessment plot–a visualisation tool to assist class enumeration in group-based trajectory modelling. *Statistical methods in medical research*, 26(5):2424–2436.

Lamoreaux, N. R., Levenstein, M., and Sokoloff, K. L. (2007). Do innovative regions inevitably decline? lessons from Cleveland's experience in the 1920s. In *Business History Conference. Business and Economic History On-line: Papers Presented at the BHC Annual Meeting*, volume 5, page 1. Business History Conference.

Maddison, A. (2007). *Contours of the world economy 1-2030 AD: Essays in macro-economic history*. OUP Oxford.

Manduca, R. A. (2019). The contribution of national income inequality to regional economic divergence. *Social Forces*, 98(2):622–648.

Martin, R. (2011). Regional economic resilience, hysteresis and recessionary shocks. *Journal of economic geography*, 12(1):1–32.

Molloy, R., Smith, C. L., and Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic Perspectives*, 25(3):173–96.

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.

Nagin, D. S. (2005). *Group-based modeling of development*. Harvard University Press.

Neffke, F., Henning, M., and Boschma, R. (2011). How do regions diversify over time? industry relatedness and the development of new growth paths in regions. *Economic geography*, 87(3):237–265.

Neil, R., Sampson, R. J., and Nagin, D. S. (2021). Social change and cohort differences in group-based arrest trajectories over the last quarter-century. *Proceedings of the National Academy of Sciences*, 118(31).

Nelson, R. R. and Winter, S. G. (1982). *An evolutionary theory of economic change*. Harvard University Press.

North, D. C. (1987). Institutions, transaction costs and economic growth. *Economic inquiry*, 25(3):419–428.

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

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, 10(1):152–197.

Petach, L. (2021). Spatial keynesian policy and the decline of regional income convergence in the USA. *Cambridge Journal of Economics*, 45(3):487–510.

Petralia, S., Balland, P.-A., and Rigby, D. L. (2016). Unveiling the geography of historical patents in the united states from 1836 to 1975. *Scientific data*, 3(1):1–14.

Putnam, R. D., Leonardi, R., and Nanetti, R. Y. (1992). *Making democracy work: Civic traditions in modern Italy*. Princeton university press.

Rees, J. and Norton, R. (1979). The product cycle and the spatial decentralization of manufacturing. *Regional Studies*, 13(2):141–151.

Rodríguez-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.

Rodríguez-Pose, A. (2020). Institutions and the fortunes of territories. *Regional Science Policy & Practice*, 12(3):371–386.

Ruggles, S., Flood, S., Foster, S., Pacas, J., Schouweiler, M., and Sobek, M. (2021). IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS.

Samuelson, P. A. (2004). Where Ricardo and Mill rebut and confirm arguments of mainstream economists supporting globalization. *Journal of Economic perspectives*, 18(3):135–146.

Singh, G. K., Daus, G. P., Allender, M., Ramey, C. T., Martin, E. K., Perry, C., De Los Reyes, A. A., and Vedamuthu, I. P. (2017). Social determinants of health in the United States: Addressing major health inequality trends for the nation, 1935-2016. *International Journal of MCH and AIDS*, 6(2):139.

Sitaraman, G., Ricks, M., and Serkin, C. (2020). Regulation and the geography of inequality. *Duke LJ*, 70:1763.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1):65–94.

Spicer, J. S. (2018). Electoral systems, regional resentment and the surprising success of anglo-american populism. *Cambridge Journal of Regions, Economy and Society*, 11(1):115–141.

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

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

Turing (n.d.). Effects of normalization techniques on logistic regression in data science. Retrieved May 30, 2023, from https://www.turing.com/kb/effects-of-normalization-techniques-on-logistic-regression-in-data-science.

Wuthnow, R. (2019). *The left behind*. Princeton University Press.

A Appendix: Correlation table for key variables

Table A.1 displays a correlation matrix for key variables involved in describing groups and estimating risk factors.

Table A.1: Correlation table: Covariates in 1940 and 2019

Variables	Annual Wages	Pop.	Pop. Density	Share Manufact.	Share Agric.	Share College	Patents p.c.	Share Black	Share Dropout	Share Foreign	90/50 Ratio	Women's Particip.
1940												
Annual Wages	1.000											
Population	0.338	1.000										
Population Density	0.235	0.891	1.000									
Share Manufacturing	0.201	0.314	0.279	1.000								
Share Agriculture	-0.701	-0.350	-0.284	-0.553	1.000							
Share College	0.456	0.105	0.050	-0.193	-0.189	1.000						
Patents per cap.	0.371	0.438	0.413	0.300	-0.334	0.101	1.000					
Share Black	-0.539	-0.002	0.015	0.141	0.373	-0.274	-0.113	1.000				
Share Dropout	-0.557	-0.008	0.047	0.258	0.262	-0.527	-0.138	0.671	1.000			
Share Foreign-born	0.610	0.326	0.266	0.051	-0.302	0.168	0.262	-0.464	-0.374	1.000		
90/50 Wage Ratio	-0.758	-0.137	-0.098	-0.286	0.718	-0.265	-0.235	0.654	0.478	-0.393	1.000	
Women's Participation rate	0.155	0.347	0.281	0.511	-0.271	0.069	0.250	0.489	0.186	0.076	0.052	1.000
2019												
Annual Wages	1.000											
Population	0.552	1.000										
Population Density	0.526	0.728	1.000									
Share Manufacturing	-0.221	-0.082	-0.057	1.000								
Share Agriculture	-0.205	-0.266	-0.293	-0.298	1.000							
Share College	0.768	0.463	0.445	-0.195	-0.184	1.000						
Patents per capita	0.580	0.294	0.244	0.065	-0.165	0.489	1.000					
Share Black	-0.037	0.138	0.168	0.075	-0.343	-0.024	-0.081	1.000				
Share Dropout	-0.336	0.004	-0.060	-0.129	0.074	-0.506	-0.188	0.184	1.000			
Share Foreign-born	0.400	0.499	0.436	-0.328	0.119	0.190	0.281	-0.056	0.426	1.000		
90/50 Wage Ratio	0.320	0.331	0.269	-0.339	-0.196	0.157	0.146	0.309	0.358	0.384	1.000	
Women's Participation rate	0.407	0.099	0.113	0.082	0.244	0.524	0.230	-0.292	-0.608	-0.005	-0.304	1.000

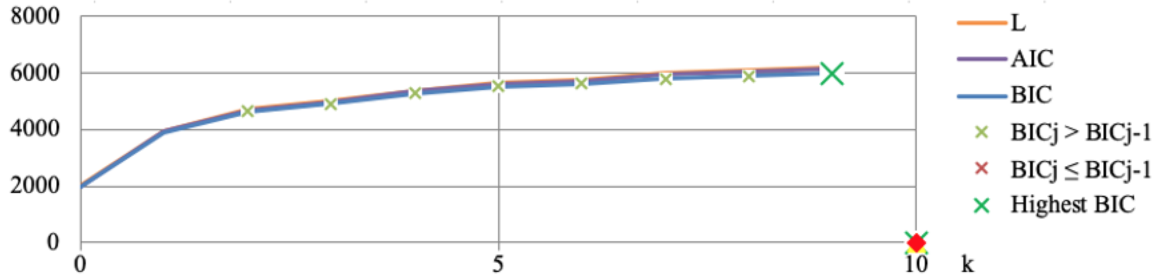
Note: Population figures are aggregates of county-level information from the Census Bureau. Population density measures are estimates based on 2010 TIGER/Geographic Identification Code Scheme (TIGER/GICS) land area information. See https://www.census.gov/quickfacts/fact/note/US/LND110210 for raw data. Other variables estimated from Census microdata. Details in the body of SI.

B Appendix: GBTM Model Selection and Diagnostics

We initially determine the optimal number of groups to include in the model; we then explore sensitivity to different polynomial orders for different trajectory groups.[16] Aiming at class enumeration, we estimated iterations of a basic model with no predictors, with the number of groups indexed by J, where $J = \{1, 2, ...10\}$, where 1 indicates the absence of latent groups in the distribution of income trajectories, and 10 pointing to the existence of ten latent groups. The upper threshold is in principle arbitrary, but a solution with more than ten groups would be impractical to work with, and risks overfitting. Initial models were run using a predetermined cubic polynomial order, chosen on the basis that it permits sufficient flexibility. The Bayesian Information Criterion (BIC) is used as the primary method of identifying J. Broadly, higher BIC values indicate greater goodness-of-fit, though as Nagin (2005) advises, this must be balanced against expert subject knowledge.

Figure B.1 displays the Bayesian Information Criterion (BIC), alongside Aikake Information Criterion (AIC) and maximum likelihood (L). Together, these indicators tell a unified story. Across each, fit improves as we add more groups to the model, up to a maximum of nine. However, while there are tangible gains in terms of fit up to five groups, after a six-group solution, improvements are at best marginal. Consequently, either five- or six-group solutions appear most robust, with only minor statistical differences between them. To adjudicate across these two competing solutions, we weighed the benefits of parsimony against substantive advantages to the six-group solution. Specifically, the six-group solution contains a group of locations, like Austin, TX, that from modest beginnings have pulled ahead over the study period. In a five-group solution, these locations are folded into a trajectory that reverts to the mean. Closer inspection reveals that this category conflates two sets of regions evolving in opposite directions. Given potential substantive interest in differentiating between these groups, we settle on a six-group solution.

Figure B.1: Trajectory fit statistics across different group solutions, ranging from 1 to 10 groups.



Note: This plot created using Klijn et al. (2017), with guidance provided by Valeria Lima-Passos.

To complete the model selection process, we iterated across all possible polynomial orders that could govern the shape of each of the six trajectory groups, considering up to a quartic polnomial.[17] Judged in terms of BIC, this yielded an optimal solution in which the shape of all but Group 1 was structured by a quartic polynomial, with Group 1 governed by a cubic

[16]Estimation was performed in Stata using *traj* - a program documented in Jones and Nagin (2013). Some diagnostics were generated using F-CAP, in R (see Klijn et al. (2017). Valeria Lima-Passos provided significant assistance with the use of F-CAP.

[17]Thanks to Jan Helmdag for providing baseline code to support this iteration process.

Table B.1: Diagnostics of trajectory group assignment accuracy, 6-group solution, polynomial order = 3, 4, 4, 4, 4, 4

Number	Group	$\tilde{\pi}$	P_j	Count	$AvePP_j$	Occ_j
(1)	Major catch up	0.136	0.138	100	0.929	82.6
(2)	Pulling ahead	0.094	0.092	67	0.929	126.9
(3)	Middle of the road	0.400	0.400	290	0.966	42.3
(4)	Mild decline	0.213	0.215	156	0.947	65.5
(5)	Fall from grace	0.121	0.121	88	0.967	216.0
(6)	Superstars	0.035	0.034	25	0.997	8567.4

Note: $\tilde{\pi}$ captures the probability of a randomly selected location being assigned to each group; P_j indicates group assignment based on the maximum posterior probability rule, while Count tallies the number of locations in each group on this basis; $AvePP_j$ is the average posterior probability for the group assigned; and Occ_j is the odds of correct classification. For further details on these diagnostics, consult the text and Nagin (2005), Chapter 5.

polynomial. Results are not materially different if cubic polynomials are used throughout.

Having settled on $J = 6$ for reasons that combine formal goodness-of-fit measures as well as substantive expertise, we carry out a series of model diagnostics.

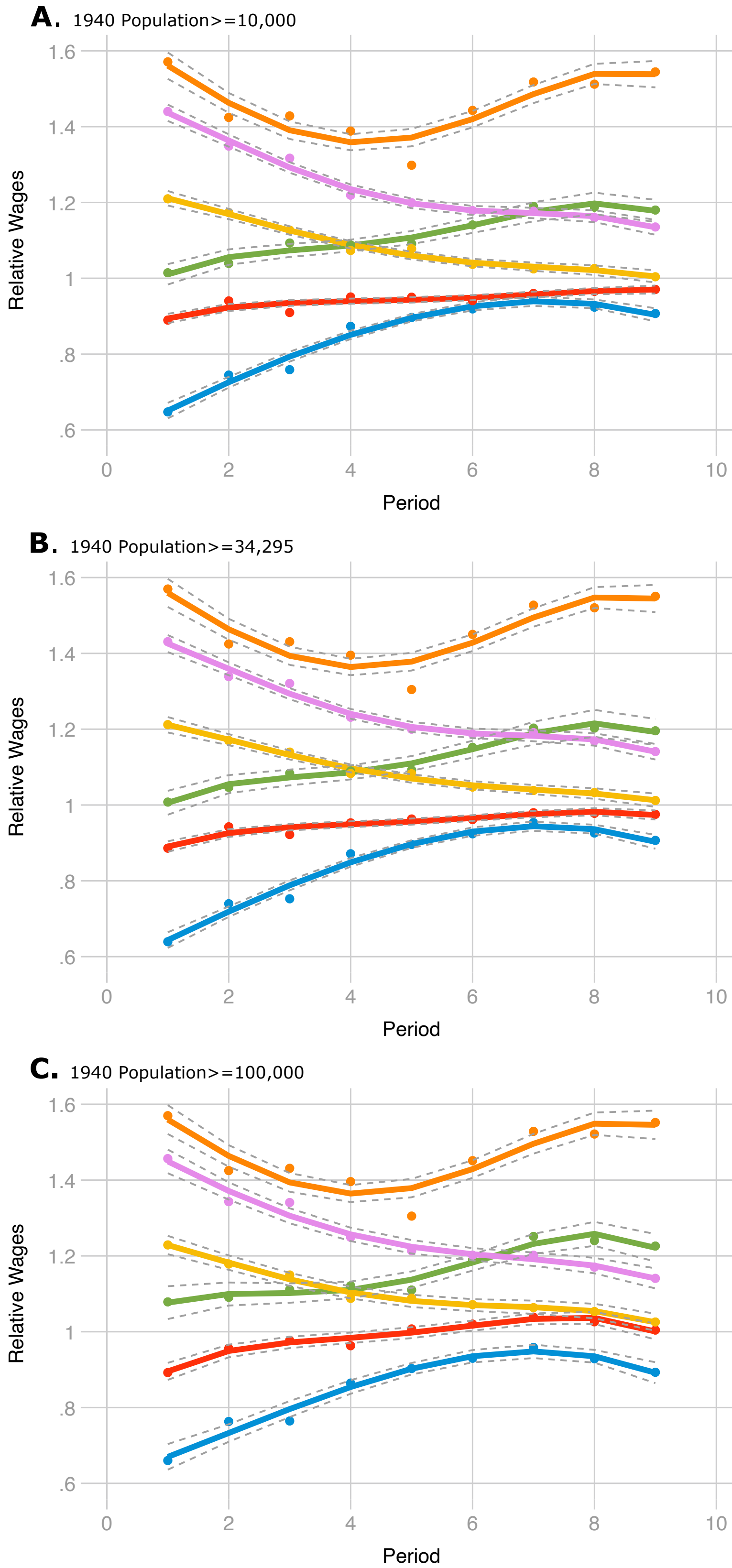
First, concerned with the possibility that these trajectories are strongly shaped by the presence of a large number of small regions, we iterated through the estimation procedure across several restrictive samples. Specifically, we explored results obtained by dropping locations whose 1940 population was below 10,000, which retained 671 out of 729 locations; dropping those below the 25th percentile in the 1940 population distribution – effectively a threshold of 34,295, which reduces the analytical sample to 545; and more extremely, dropping all locations with 1940 populations of less than 100,000, which retains an analytical sample of just 297 cities. Resulting trajectory plots are shown in Fig B.2. In each case, these plots strongly resemble those shown in Figure 4, suggesting the strength of the chosen model.

In Table B.1, we report a series of more formal model diagnostics, described in detail in Nagin (2005). These measures broadly confirm the success with which a six-group solution accurately assigns locations to groups. The column titled $\tilde{\pi}$ indicates the probability that a randomly selected location will belong to a particular group, producing estimates that have already been reported in Figure 4.

P_j measures the proportion of the sample assigned to each group-based on a different metric: maximizing the posterior group membership probability. In this procedure, calculated separately after estimation, each location receives an estimate of the probability that they belong to each group, contingent on its observed behavior over time. Locations are then assigned to a single group-based upon the location for which it is mostly likely to belong. P_j indicates the proportional distribution of regions to groups on this basis. As Nagin (2005) suggests, in an imagined situation in which we are entirely certain to which group each unit belongs, P_j and $\tilde{\pi}$ will be identical. Given the high degree of observed consistency between these two proportions in Table B.1, we can be confident that assignment accuracy is high. The 'Count' column reports the number of locations allocated to each group on the basis of maximum posterior probability.

Another useful measure for model evaluation is $AvePP_j$, or the average maximum posterior probability. Each location receives an estimate of the probability that they belong to each group, and is assigned to the single group for which its probability is highest. For each resulting

Figure B.2: Group-based trajectories estimated on subsamples determined by initial population.



Note: N=671 for Panel A; N=545 for Panel B; N=297 for Panel C. Threshold for Panel B is the 24th percentile of population in 1940.

group j, $AvePP$ is calculated by taking the average of these likelihoods among its 'members.' A group with a value of $AvePP$ equal to one would be one in which each of its members was assigned to that group with perfect certainty. For a hypothetical group with an $AvePP$ equal to 0.5, we would be as confident about its average location's group membership as we are about a coin toss. In actuality, the groups in our preferred mixture model receive $AvePP$ that range from a low of 0.929 to a high of 0.997. These very high probabilities indicate that groups are assigned with little likelihood of misclassification.

Finally Occ_j indicates the odds of correct classification, yet another application of the posterior probabilities in the service of gauging the effectiveness with which units are classified to groups. Higher values indicate that the resulting classification of locations to groups performs better than random assignment, with values above 5 considered as a lower bound on adequacy. As per Nagin (2005), Occ_j is given by the formula:

$$\frac{AvePP_j/1 - AvePP_j}{\tilde{\pi}_j/1 - \tilde{\pi}_j} \tag{4}$$

with the numerator equivalent to the odds that a location is correctly versus incorrectly classified, divided by similar odds based on random classification. The group with the lowest Occ_j remains far above that threshold, while for other groups this hurdle is cleared dramatically. Overall then, the diagnostics presented in Table B.1 highlight the high degree of certainty that locations are well allocated using our favored, six-group trajectory model.

B.1 Describing groups

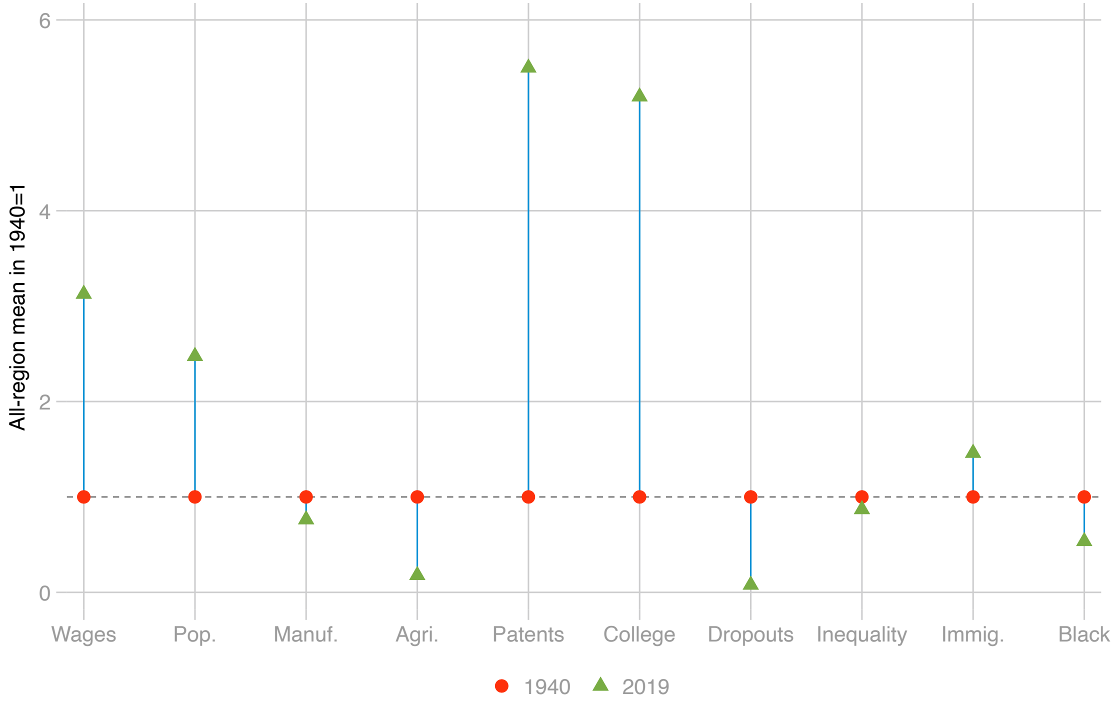
The main text describes groups on the basis of the evolution of changes in urbanization, economic and social structure. Since that description is undertaken in reference to national trends, here, in Figure B.3, we present national changes in these indicators on their own.

Over the 79-year study period, inflation-adjusted wages grow nationally by a factor of just over three, while population grows only slightly less. The share of employment in manufacturing and agriculture declines, the latter very sharply. Meanwhile, rates of patenting per capita and the share of workers with at least four years of college both grow dramatically. The high school dropout rate declines very sharply. Income inequality remains fairly constant, as measured by the ratio of annual wages at the 90th to the 50th percentile. Immigration grows, while the share of blacks in total employment declines.

B.2 Modeling Risk Factors

As a means to specify the role of urbanization, economic and social structure in shaping locations' membership in a particular trajectory, we turn to multinomial logistic regression, estimated simultaneously with the trajectories to directly incorporate assignment uncertainty. This regression aims to determine whether, controlling for other potentially relevant selection factors, differences in initial conditions for a particular endowment affect a location's probability of belonging to group x versus group y. As distinct from simple between-group univariate comparisons of the kind discussed in the 'trajectory group profiles' section of the main paper, regression results described here specify a relationship between π_j and our set of location-specific

Figure B.3: National trends in urbanization, economic structure, social structure, and wages.



Note: Values in 2019 represent deviations against initial (1940) values. 'Wages' is annual wage and salary income; 'Pop' represents population; 'Manuf' is the share in employment in manufacturing setcors; 'Agri' measures share of employment in agricultural industries; 'Patents' is granted patents per capita; 'College' is the share in employment with at least four years of college education; 'Dropout' is the share in employment with less than a high school diploma; 'Ineq' in the ratio of income at the 90th percentile to the 50th percentile; 'Immig' is the share in employment that is foreign-born; and 'Black' is the share in employment who self-identify as black. Data are described in more detail in the SI text.

covariates described above. They also permit us to consider if certain predictors that appear to be important in univariate profiles remain so after accounting for other relationships of interest.

Since we cannot feasibly compare each group to each other group, in this initial piece of work we focus on contrasts drawn against the group that most closely resembles the national average: Middle of the Road. Not only is this group 'typical' in this sense, it is also the group with the largest membership in terms of number of locations, measured using the maximum posterior probability rule.

Table B.2 presents coefficients and standard errors for predictors for various measures of urbanization, economic structure and social structure, as described in the text. Parameter estimates for groups for this model are not separately reported, but are nearly identical to the base model with no covariates, described in Figure 3 in the main paper. Positive coefficients indicate that higher values of the predictor are associated with a greater likelihood that a location belongs to the group at hand as opposed to the reference group. Negative coefficients imply the opposite relationship. Hence, the coefficient -1.97 for the population variable for the Catch Up group indicates that locations with higher population levels in 1940 are more likely to be in the Middle of the Road group as opposed to the Catch Up group. This relationship is deemed to be significant at a threshold $p < 0.001$. Note that this model does not permit us to, for instance, distinguish formally between the extent to which initial patenting levels predict

Table B.2: Predictors of group membership in income trajectories

	Catch Up	Mild Decline	Pulling Ahead	Fall from Grace	Superstars
Reference group: Middle of the Road					
Population	-1.97***	0.99**	1.27***	2.88***	3.00 ***
	(0.59)	(0.31)	(0.28)	(0.44)	(0.44)
Manufacturing	0.08	-0.03	-0.08**	-0.09*	-0.15 **
	(0.05)	(0.03)	(0.03)	(0.03)	(0.05)
Agriculture	0.35***	-0.25***	-0.07	-0.26***	-0.34**
	(0.08)	(0.05)	(0.05)	(0.08)	(0.13)
Patenting	-0.39*	0.03	0.02	0.07	0.09*
	(0.18)	(0.03)	(0.03)	(0.04)	(0.05)
Graduates	-0.05	0.30*	0.00	0.44*	0.07
	(0.14)	(0.13)	(0.12)	(0.20)	(0.25)
Dropouts	0.05	0.06	-0.02	-0.14	-0.32**
	(0.04)	(0.04)	(0.04)	(0.07)	(0.10)
Inequality	2.89***	-7.26***	-3.72***	-22.51***	-20.30***
	(0.78)	(1.09)	(0.76)	(3.12)	(3.76)
Foreign-born	-0.30***	0.25***	0.22***	0.39***	0.57***
	(0.08)	(0.06)	(0.06)	(0.08)	(0.10)
Black	0.01	-0.12*	0.11***	0.10	0.25**
	(0.03)	(0.05)	(0.02)	(0.09)	(0.09)

Note: Coefficients are presented in the top row, along with significance stars, where * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors are included in parentheses, rounded to the nearest tenth. Population is measured in 100,000s; Manufacturing, Agriculture, Graduates, Dropouts, Foreign-born and Black are shares of total employment (scaled 0-100); Patenting is measured as the number of granted USPTO patents per 100,000 population; Inequality is the ratio of the local 90th percentile of income to 50th. All predictors in the model are set to 1940 values.

whether a location is likely to follow the Superstar or Fall from Grace trajectories – though tests of the equality of any two coefficients can be estimated in the manner of Clogg et al. (1995).

Odds ratios are presented in the main text instead of raw coefficients, as the former permit more straightforward measurement of the strength of relationships under investigation. The odds ratio represents the extent of the change in expected probability of belonging to group j versus the reference group that is associated with a one unit change in the underlying predictor.

C Appendix: Income or Population?

The reason we have investigated trajectories of income in this research is that they are meaningful indicators of dynamic spatial development inequalities. The use of nominal income in the international convergence-divergence tradition has been relatively unproblematic because of the large gaps in international prices, with PPP conversions modifying, but not generally calling into question any picture of development hierarchies and clubs. Though other measures have been explored (Jones and Klenow, 2016), comparisons are principally made using per capita gross domestic product or income, often adjusted for differences in purchasing power (Maddison, 2007). At the subnational scale investigated in this paper, these issues are also relevant and widely examined (Carlino, 1992; Drennan et al., 1996; Gaubert et al., 2021), but with some additional dimensions to take into account. Just as in international comparative development research, between regions there can be important differences in living costs and quality of life. At the same time, internal migration is much less costly and much more prevalent than across international borders. Such distinctive features of the inter-regional as opposed to international context require us to interpret carefully the meaning of income differences as indicators of the sub-national development disparities.

In some versions of the standard spatial equilibrium framework households enjoy relatively frictionless mobility opportunities, and choose regions by arbitraging a wide variety of preferences, with the key ones consisting of nominal income, housing type and cost, a range of priced and unpriced amenities, and the avoidance of dis-amenities Glaeser (2008). This leads to the conclusion that real income – and even more so, total real utility – variations will converge to such an extent that the most robust indicator of regional development is population change.

Before continuing this discussion, Figure C.1 shows the changing population shares of our six groups of income dynamics. The shares of the two groups with persistently high income (Superstars) or rapidly growing income ranks (Pulling Ahead) have increased, in contrast to all the other groups. We cannot determine whether agglomeration of high-wage activities is the principal cause of population growth, or the reverse, whether income growth is due to differential population sorting, and in any case, the two channels of change are mutually compatible. This gives a picture of the already-skilled places or the rapidly increasingly skilled places becoming bigger in population terms relative to all other places.

To return to the spatial equilibrium framework, the assumption of real income or total utility convergence is not supported by available evidence. At the inter-regional scale, real incomes show both nominal and real income divergence. Though high housing costs do erode high nominal incomes in some cities, studies have not found clear evidence of equalization (c.f. Moretti, 2013; Kemeny and Storper, 2012, 2020; Diamond and Moretti, 2021).[18] In any case, in this study, we have carefully adjusted nominal incomes for local consumer prices and housing costs, leading to a reasonable approximation of real incomes over time. The claim that real utilities are equalizing is a more ambitious claim of spatial equilibrium theories, and underlies the assertion, in some of them, that population change is the only robust indicator of regional development dynamics. But empirical research on utilities is still quite limited. There is a growing body of literature, cited in the introduction and literature review to this article, that shows – on many fronts – that amenities and quality of life differ very significantly between regions, enough to strongly imply that there it is not reasonable to assume anything like total utility equalization. Even in the mainstream tradition, Diamond (2016) recognizes that utility preferences are not only non-homothetic, and that that nominal income differences have now reached a point where some groups can aspire to certain things that other groups cannot, with non-overlapping income ability to spatially access certain kinds of amenities, and that amenities are strongly endogenous to regional incomes. Thus, though amenities are varied and difficult to rank in a manner that accounts for heterogeneous preferences, larger, higher-income cities play host to many of them, likely as an outgrowth from those same high incomes (Diamond, 2016). Finally, and crucially, we know that a major slowdown of internal migration is

[18]This could still be happening on the margin.

limiting the flows of certain people moving to higher-income and amenity locations (Molloy et al., 2011; Ganong and Shoag, 2017; Kaplan and Schulhofer-Wohl, 2017). While model precepts emphasize identifying the properties of an urban system in equilibrium or deviations from it as mere frictions, the current consensus is that nominal spatial income disparities and associated interaction effects with spatially concentrated amenities represent a substantial problem to be understood and addressed (Ganong and Shoag, 2017; Austin et al., 2018; Gaubert et al., 2021).

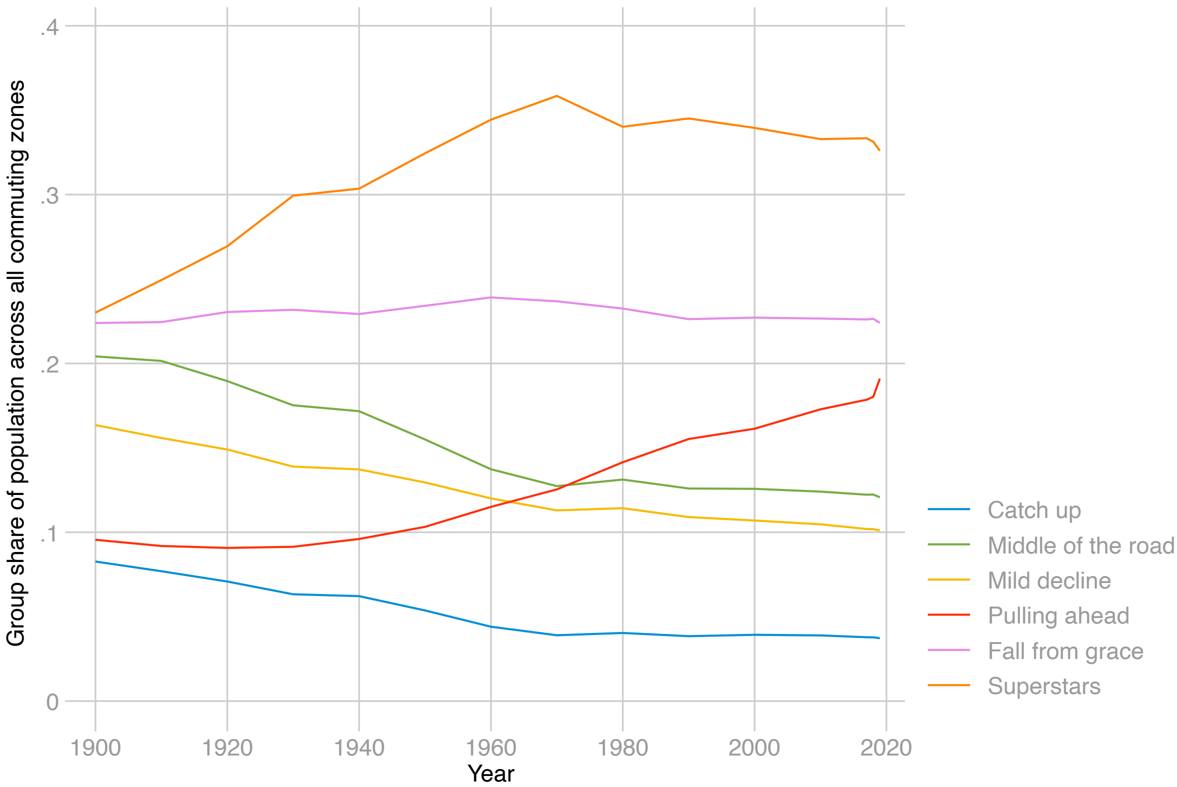


Figure C.1: Group shares of total population, 1900-2019

Note: Each line represents changes in the ratio of the total population across U.S. commuting zones for a particular group to the total population across all U.S. commuting zones. Hence, a value of 0.2 means that, in a given year, a particular trajectory group makes up 20% of the all-locations population. See main analysis for the derivation of groups.