NBER WORKING PAPER SERIES

INDUSTRY AND IDENTITY:
THE MIGRATION LINKAGE BETWEEN ECONOMIC AND CULTURAL CHANGE IN 19TH CENTURY BRITAIN

Vasiliki Fouka
Theo Serlin

Working Paper 33114
http://www.nber.org/papers/w33114

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2024

The authors would like to thank Rikhil Bhavnani, Volha Charnysh, Gary Cox, Amir Goldberg, David Laitin, Jonathan Rodden, Kenneth Scheve, Alain Schläpfer, Marco Tabellini, Yanos Zylberberg and seminar participants at Jinan University, the University of Copenhagen Workshop on Migration and Culture, the Stanford GSB Workshop on Cultural Evolution and Organizational Theory, the 2024 APSA Annual Meeting and the 2024 ASSA Annual Meeting for helpful feedback. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Industry and Identity: The Migration Linkage Between Economic and Cultural Change in 19th Century Britain
Vasiliki Fouka and Theo Serlin
NBER Working Paper No. 33114
November 2024
JEL No. J6, N0, N33, N63, Z1

ABSTRACT

How does economic modernization affect group identity? Modernization theory emphasizes how labor migration led to the adoption of common identities. Yet economic development may reduce incentives to emigrate, preserving local cultures. We study England and Wales during the Second Industrial Revolution, a period characterized by the development of new industries and declines in transportation and communication costs. Using microdata on individuals' names and migration decisions, we quantify identity change and its variation across space. We develop and estimate a quantitative spatial model in which migration and cultural identities are inter-dependent. Different components of economic modernization had different effects on identity change. Falling migration costs homogenized peripheral regions. In contrast, industrial development led to heterogeneity, increasing the overall prevalence of the culture of London, while also creating local identity holdouts by reducing out-migration from industrializing peripheries. Modernization promotes both national identities and persistent local identities in peripheral regions that industrialize.

Vasiliki Fouka
Stanford University
Department of Political Science
616 Jane Stanford Way
Stanford, CA 94305
and CEPR
and also NBER
vfouka@stanford.edu

Theo Serlin
Princeton University
404A Robertson Hall
Princeton, NJ 08540
tserlin@princeton.edu

A data appendix is available at http://www.nber.org/data-appendix/w33114

1 INTRODUCTION

A long line of thought links the development of industry to the reconfiguration of individuals' cultural ties and the emergence of new group loyalties. From the Communist Manifesto to classical modernization theory, scholars have seen industrialization—and the host of developments accompanying it, such as education, urbanization and the spread of mass media—as a force that promoted the breakup of local and particularistic cultural affiliations (Marx and Engels, 1978 [1848]) and their consolidation into overarching identities that led to the formation of modern nations (Gellner, 2006 [1983]).

Most of this literature emphasizes migration as one of the central mechanisms connecting industrial development and cultural change. The emergence of industrial centers attracted migrants from agricultural regions both near and far away, who in turn required a common language and set of cultural norms in order to communicate. An inflow of migrants from diverse cultures into rapidly industrializing locations was seen as favoring the adoption of a single identity that served as a coordinating device (Gellner, 2006 [1983], p.59; Weber, 1976, p.78).

Yet theoretically, the implications of migration flows for the cultural landscape that follows industrial change are not straightforward. Industrializing areas may experience higher in-migration, a force that pushes towards cultural change. At the same time, industrial development should lower out-migration, a force in the direction of cultural retention. Crucially, how these two opposing forces offset each other in a single location depends on what happens everywhere else. The geographic location of one industrial center relative to others affects where in-migrants will be drawn from and whether they will mainly originate from different cultures promoting cultural change or from the local culture favoring local identity preservation. Seminal theories of industrialization and identity formation that recognize a role for migration do not draw out its full theoretical implications.

In this paper we attempt to fill this gap. We explore theoretically and empirically the role of industrialization-induced migration for cultural change in the context of the world's first industrializing economy: England and Wales during the Second Industrial Revolution. Between 1851 and 1911, manufacturing growth increased dramatically and the spatial pattern of industrialization shifted away from Lancashire and textile production to new hubs in the Northeast and South Wales, where the presence of coalfields attracted metallurgical industries. At the same time, the development of the railway network and breakthroughs in communication technologies such as the telegram reduced the costs of migration by increasing connectivity and facilitating communication. We use rich census microdata on households' migration and cultural choices to study how these changes affected patterns of internal mobility and, through them, the resulting cultural map. An advantage of studying the industrialization-migration-culture nexus in the case of Britain, as opposed to other well-known cases of identity transformation such as France (Weber, 1976), is that Britain did not engage in state-led nation-building efforts. There was no conscription and education was administered at the local level. This case allows us to examine how identity reacts to economic changes in the absence of central planning.

We begin by documenting the pattern of cultural change in England and Wales between 1851 and 1911. To quantify the landscape of local cultures on the eve of the Second Industrial Revolution, we rely on detailed individual-level data from the 1851 census. We use surnames of household heads born before 1800 to identify a stable component of local culture. A spectral clustering algorithm applied to surnames yields nine cultural clusters. These clusters are meaningful representations of local cultures: they trace the boundaries of identifiable regions, such as East Anglia and the West Country, and are reflected in rates of migration and intermarriage, which are significantly higher within clusters than across them.

We then use first names—a faster-moving outcome than surnames—to examine how identification with each of these local cultures changed between 1851 and 1911. We assign first names to cultural clusters by computing an empirical index of distinctiveness (Fryer Jr and Levitt, 2004; Fouka, 2019) and then compare the cultural content of names of people born 1851–1860 to those born 1901–1910. We document three patterns. First, between 1851 and 1911, England's cultural landscape underwent significant homogenization (Fact 1). Entire cultural regions disappeared from the map and names that increased in popularity were most associated with the cultural cluster of the Southeast of England, which includes London. Second, homogenization displayed a center-periphery gradient, with peripheral regions experiencing stronger cultural convergence towards the Southeast and more rapid declines in their local cultures (Fact 2). Third, the association between industrialization, proxied by the presence of coalfields, and local cultures also varied depending on geographic location (Fact 3). In the center of the country, the presence of coal was associated with declines in local cultures. Instead, no such relationship existed in the peripheries, where coal deposits created relative cultural holdouts.

We provide descriptive evidence that these patterns were plausibly mediated by migration. During the period under study internal mobility increased, with out-migration rates from some regions rising up to 15 percentage points. People left agricultural areas and moved towards regions with high manufacturing growth. In turn, both in- and out-migration were associated with changes in the prevalence of the Southeastern and home cultures. Migration and cultural choices were also interrelated at the individual level. Having a name indicative of a particular cultural cluster increased the probability of migration to that cluster, suggesting that people were more likely to migrate to culturally similar areas. In turn, parents tended to give children first names that were characteristic of likely future migration destinations, as predicted by the location of coal deposits. This indicates that naming choices anticipated migration decisions.

We build on these empirical observations to develop a theoretical model, in which agents choose where to migrate based on economic returns and their preference to be with others of the same culture. Cultural choices are made by agents' parents, who have tastes for specific cultures, but also wish to maximize their offspring's future expected welfare. This model allows us to study in a principled way the equilibrium effects of industrialization on local identities, accounting for interdependence across regions and offsetting changes in in- and out-migration in response to economic change. Additionally, this approach allows us to separate what are conceptually distinct components of modernization—industrial activity on the one hand, and changes such as connectivity and transport costs on the other—and examine their distinct contributions to the identity changes we observe.

The model yields equilibrium culture-origin-destination migration flows and cultural choices at the origin, quantities which we observe in census data. Using an instrumental variables strategy based on the interaction between geographic locations, historical cultures, and coal deposits, we estimate two theoretical parameters: the "homophily elasticity," which captures how migration decisions respond to the number of people from the same culture present in a destination, and the "culture elasticity," which captures how the choice of culture responds to the expected benefits of migration under that culture. Using the estimated elasticities we can then back out unobserved taste parameters for specific destinations among migrants and specific cultures among parents. We validate the model by showing that these estimated parameters correlate strongly with measures of cultural proximity, such as surname similarity and religious distance.

We use the model to conduct counterfactual exercises under different values of economic fundamentals. We first show that the model broadly rationalizes the empirical patterns in the data. The observed change in the choices of first names associated with different cultural clusters between 1851 and 1911 strongly correlates with the model's "prediction" based on the difference between 1911 cultural choices and a counterfactual estimated under 1851 parameters (holding fixed non-economic factors influencing culture). This suggests that, even though our model focuses on a specific mechanism and abstracts from many undoubtedly relevant developments for identity change between 1851 and 1911, it nonetheless explains a significant part of the change that occured during the period. Next, we turn to the two main exogenous parameters of the model: the real wage (industrialization across regions) and migration costs. By counterfactually setting each in turn to its 1851 value, we can distinguish its contribution to observed cultural changes.

The counterfactuals match the stylized facts we document. The main force responsible for the rise in the culture of the Southeast (Fact 1) is industrial development reflected in the real wage. Two channels drive this result. First, the Southeast of England experienced an increase in manufacturing growth and real wages during this period, decreasing out-migration from that region. Second, real wage increases also occurred in other regions which were well-situated to attract migrants from the Southeast, increasing the prevalence of the Southeast culture elsewhere in the country. While there may have been other historical or political factors outside the model driving the prominence of the Southeast as a focal identity, these channels provide an explanation for why the culture of London dominated England and Wales by 1911 relying solely on the interplay of industrial development and migration pull factors.

Instead, the core-periphery gradient in cultural homogenization (Fact 2) was mainly driven by falling migration costs, which facilitated migration over longer distances. While the decrease in migration costs helped spread the culture of the Southeast to the peripheries, industrial development moderated this effect. Peripheries with coal deposits were able to resist the tendency of other peripheries towards more rapid cultural loss (Fact 3). This is once again a result of the offsetting forces triggered by industrial growth on in- and out-migration. Compared to the center of the country, in the periphery the lower rates of migration to other cultural clusters caused by the presence of coal were offset to a much smaller degree by in-migration from the outside. This helped preserve local cultures in regions like Wales and the Northeast of England.

We also use the model to answer the broader question of when local industrialization supports or undermines local culture. While the counterfactual exercises yield equilibrium cultural choices in all locations, this analysis examines how cultural choice in each individual location changes in response to changes in the real wage in that location, holding other parameters constant. This exercise then allows us to quantify the heterogeneous effects of industrialization. In peripheral locations, industrialization promotes cultural retention by reducing the need to out-migrate. In more central areas, economic opportunity attracts migrants from other districts, leading to loss of the local culture.

Our study makes five main contributions. First, we contribute to both twentieth century and more recent literatures linking industrialization to cultural change. Seminal scholarship tends to view industrialization as a homogenizing force (Marx and Engels, 1978 [1848]; Deutsch, 1966; Gellner, 2006 [1983]), though some work has also pointed out that uneven industrial development can drive differentiation and strengthen pre-existing cultural cleavages (Gellner 1964, 166–168; Deutsch 1969, 20–25; Hechter 1977). More recent work like Green (2019) and Green (2022) also links industrialization to assimilation into larger cultures. Similar to Green (2022), our framework also focuses on economic incentives and bottom-up processes of assimilation rather than state-led nation-building. Yet by modeling migration and cultural choice in a spatial equilibrium framework we show that the effects of industrial development can be heterogeneous and the identities that dominate depend on the spatial distribution of economic activity and local cultures. Our findings suggest that the retention of local identities in peripheral locations can also be generated by mechanisms other than economic domination (Hechter, 1977) or status considerations and elite incentives (Gellner, 1964).

Second, and more broadly, we contribute to a historical and sociological literature on identity change in Europe during the 19th and early 20th century. Classical modernization theory posits many forces that drove the formation of unifying identities in Europe, such as education (Gellner, 1964), cultural diffusion through print media (Anderson, 1983), or the construction of infrastructure networks across markets (Weber, 1976), but has also emphasized the role of internal mobility (Deutsch, 1969; Weber, 1976; Gellner, 2006 [1983]). We model this mechanism explicitly in a spatial equilibrium framework. In so doing, we empirically validate insights in the modernization literature on the links between industrialization, migration and cultural convergence. At the same time, we also show that migration can act as a force limiting convergence and preserving peripheral identities.[1] Our model also allows us to separate industrial development and falling costs of migration, components of economic modernization that are often bundled together.

Third, we contribute to the literature on the determinants of cultural choice. Much work across disciplines considers identity and culture to be an individual decision responding to economic and non-economic incentives (Waters, 1990; Laitin, 1995; Lazear, 1999). We develop a model in which economic incentives influence cultural choices by affecting where people migrate, and whether they seek to coordinate on culture with the people already there. The model combines elements of the framework of Bisin and Verdier (2001), in which parents choose their offspring's culture based on expected returns and their own preferences, with insights from Laitin (1994), who views the choice of culture as a coordination game (see also Laitin 2007). Our quantitative model allows us to quantify the importance of these economic and coordination forces.

Fourth, we apply empirical methods from a growing body of work on quantitative spatial models to study how industrialization affected cultural homogenization in England in the second half of the 19th century. Recent work in urban and international economics has

[1] Second-generation modernization literature that focused on Africa also proposed that modernization may strengthen, instead of eradicating, narrow group identities such as ethnicity (Bates, 1974), with mixed empirical evidence (Robinson, 2014; Pengl, Roessler and Rueda, 2022). This strand of research examines a set of mechanisms distinct from those we focus on and more related to group competition over the gains of modernization.

developed tractable models of economic geography (Allen and Arkolakis, 2014; Redding and Sturm, 2008). These models can rationalize observed migration and commuting flows, and, when calibrated to observed data, generate accurate predictions of how changes in trade costs alter the spatial distribution of economic activity (Redding and Sturm, 2008; Ahlfeldt et al., 2015). Beyond applying this methodology to the study of cultural change, our analysis connects to this literature in two ways. Monte, Redding and Rossi-Hansberg (2018) and Caliendo et al. (2018) show that the spatial structure of the economy generates heterogeneity in the impact of shocks to labor demand and productivity on employment and output. We show that the same is true of the relationship between economic growth and cultural choices. Bryan and Morten (2019) and Morten and Oliveira (2023) study how migration costs affect economic output by influencing whether workers are allocated to the regions in which they will be most productive. We show that preferences for cultural sorting also affect the allocation of workers across regions and thus also influence total output.

Fifth, and more broadly, we contribute to a quantitative literature studying the determinants of cultural convergence across space. Most studies focus on the role of trade and globalization in influencing cultural diversity (Maystre et al. 2014; Olivier, Thoenig and Verdier 2008; see also Bisin and Verdier 2014 for a review). We study instead the role of migration. Related to our setup, Rapoport, Sardoschau and Silve (2020) also examine how migration contributes to cultural convergence allowing a role for homophily in dictating migration decisions. Different from that study, we take a structural approach to studying patterns of cultural homogenization within a country in response to changes in economic fundamentals.

The rest of the paper proceeds as follows. Section 2 provides a background on economic modernization in Britain during the Second Industrial Revolution. In Section 3 we describe the data and the methods we use to quantify early local identities and cultural choice and present descriptive patterns of identity change in England and Wales. Section 4 provides descriptive evidence on the role of internal migration in driving identity change and on the link between migration and identity choice. Section 5 introduces our theoretical model, while Section 6 explains how we use the observed data to estimate the model and validate recovered structural parameters. In Section 7, we conduct counterfactual exercises to decompose the contributions of different forms of economic change to observed cultural change. In Section 8 we use the model to understand how local economic development influences cultural choices. Section 9 presents additional extensions and robustness checks, addressing, among others, how factors like transportation infrastructure, local wage differentials and emigration influence the results of the model. Section 10 concludes with implications of our study for identity formation and nation-building.

2 ECONOMIC CHANGE IN ENGLAND AND WALES DURING THE SECOND INDUSTRIAL REVOLUTION

During the second half of the 19th century Britain underwent rapid economic transformation. This was marked by two major developments. The first was a dramatic change in the sectors and spatial patterns of economic activity. In 1851, British manufacturing was dominated by textiles, the staple industry of the First Industrial Revolution. Over the period 1851–1911, the steel, chemicals, and engineering and secondary metals industries came to rival the textiles industry. These industries developed in response to technological developments: the Bessemer Process for making steel (1856), the Solvay Process for making ammonia (1861), and the replacement of wooden sailing ships with steel steamships in the 1870s (Crouzet, 1982). Wool manufacturing made up 9% of manufacturing employment in 1851, but 4% in 1911; machinery manufacturing increased from 3% to 10% of manufacturing employment.[2]

The growth of new industries altered the spatial distribution of economic activity. Figure 1 shows the share of employment in manufacturing by registration district in 1851, and the change in the log number of manufacturing workers between 1851 and 1911. In 1851, manufacturing was concentrated in Lancashire, the center of the cotton industry. Between 1851 and 1911, rural areas in the Southwest, East Anglia, Wales, and the North declined and already industrialized parts of Lancashire and the Midlands experienced continued growth. Growth did not just conform to existing patterns of development. The new steel and metals industries were located near major coalfields in South Wales and the Northeast. During this period London and the Southeast also experienced rapid growth, in part due to new industries locating close to investors in the City of London (see also Geary and Stark 2015).

A second major development of the late 19th century was a dramatic reduction in the costs of mobility. Part of this was driven by the spread of technologies like the railway and telegraph, which greatly facilitated migration across regions. The first public railway using steam-powered locomotives that carried both passengers and freight was built in 1830 and connected Liverpool to Manchester. While major towns were connected from earlier on (Bogart et al., 2022), by 1881 the railway network covered almost all of England and Wales, extending to 25,000km of lines. Passenger journeys increased from 20 to 1,300 million between 1841 and 1911 (see also Figure 2). Much of this increase, particularly between 1851 and 1911, was accounted for by third-class passenger travel, enabled by the Railway Regulation Act of 1844 which provided compulsory third-class accommodation on trains at a low price meant explicitly to facilitate travel in search of work for the poorer segment of the population.

[2] These figures were calculated from I-CeM census data (Schürer and Higgs, 2014). Figure A.1 shows employment in different manufacturing industries in 1851 and 1911.

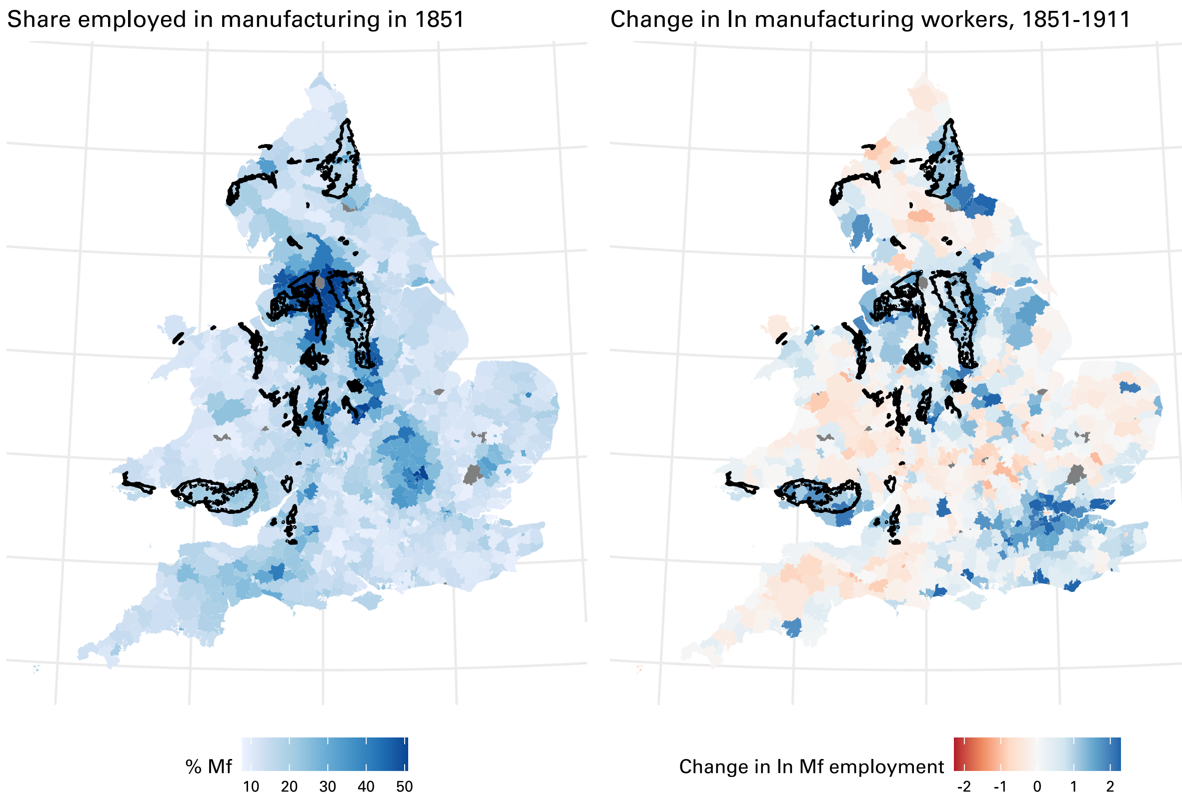

Figure 1: Industrialization in 1851, and growth over the 1851–1911 period

The map on the left shows the share of the population employed in manufacturing, the secondary sector, in 1851. The map on the right shows the change in the log number employed in the secondary sector between 1851 and 1911. In both maps, coal deposits are outlined in black.

Railways provided a means for the working class to travel en masse over large distances for the first time in British history (Shaw-Taylor et al., 2018).

The development of railways also facilitated communication across regions. In 1838, Parliament authorized the carriage of mail by railway and in 1840 the Uniform Penny Post set a uniform delivery rate of one penny for mail postage anywhere within Britain and Ireland, greatly increasing accessibility and use of the postal system (Schwartz, 2023). The per capita number of letters sent increased steadily throughout the 1851–1911 period (Figure 2). Railways also enabled the uptake of the new technology of the electric telegraph, or telegram, which was initially used as a railway signaling system before becoming a more general means of communication in the second half of the 19th century (Fava-Verde, 2018) (see also Figure 2). The reduction in communication costs promoted migration both by providing information on destinations and by reducing the cost of migrants' communication with their hometowns.

Other processes driven by economic modernization also reduced the costs of mobility. The continuing increase in aggregate industrial production pulled people from rural areas

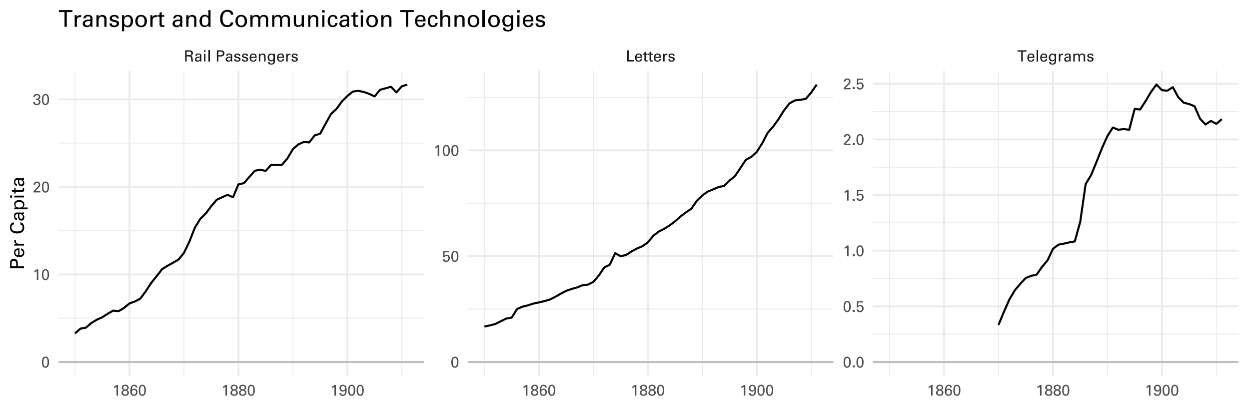

Figure 2: Expansion and Growing Accessibility of Rail, Mail and Telegrams

The figures plot rail passengers, letters, and telegrams, divided by population in the UK between 1850 and 1911. Source Palgrave Macmillan Ltd (2013).

into cities, breaking their ties to traditional social networks. Severing these links enabled further mobility. As demand for specialized labor increased due to the rise of new industrial sectors, mobile individuals would find it easier to move to locations with specific industries demanding their skills (Bryan and Morten, 2019). The period was also marked by the growth of the education sector and rising literacy, as successive education acts in 1870 and 1880 established a system of compulsory, free, primary education in England and Wales.[3] Education contributed to enabling mobility by changing people's values and attachment to their local origins and by providing information on the benefits of migration.

3 EMPIRICAL PATTERNS OF IDENTITY CHANGE

3.1 *Data*

Our main data source is the I-CeM full count microdata from the Census of England and Wales, 1851–1911 (Schürer and Higgs, 2014). Two features of this data make it especially useful for our purposes. First, we have access to full names in the 1911 census, which we use to measure cultural choices, and the surnames of household heads in earlier censuses, which we use to cluster locations into cultural groupings. Second, the census recorded granular information on birthplaces, which allow us to trace out lifetime migration.

Our unit of analysis throughout is the registration district, fixed to 1851 boundaries. Registration districts were administrative units created in the 1830s to register births, deaths and marriages.[4] The advantage of focusing on registration districts rather than smaller units

[3] Figure A.2 shows the increase in primary school pupils and teachers.

[4] In a few cases a registration district falls into more than one county (the largest administrative unit in

such as parishes is that other social and economic data from this period which we use to validate our model, for instance attendance at churches of different religious denominations, were reported at the district level. We link birthplaces to registration districts using crosswalk files from Day (2016) and a GIS from Satchell et al. (2016).

3.2 *Cultural Clusters Before the Second Industrial Revolution*

Our analysis focuses on how the economic changes brought about by the Second Industrial Revolution influenced identity change. As a prerequisite for this, we first need a way to measure local identities prior to the developments we analyze. Historically, Britain has been characterized by a high degree of diversity in local cultures and customs, driven not only by differences between its constitutive nations, but also differences within England (see for instance Homans (1969)). Variation in local geography and soil types, as well as centuries of limited communication across regions had generated "a patchwork in which uncertain areas of Welshness, Scotishness and Englishness were cut across by strong regional attachments, and scored over again by loyalties to village, town, family and landscape" (Colley, 2005, p.17). To quantify the distribution of these regional identities, we use a data-driven way to allocate districts to cultural "clusters" or regions that display high cultural similarity.

Our approach uses information on the surnames of household heads born before 1800, recorded in the 1851 census. We record the share born in each district with each surname. This data predates the Second Industrial Revolution. In 1800, the First Industrial Revolution was underway, but growth was slow—Antràs and Voth (2003) estimate productivity growth of 0.2% per annum 1770–1800—and confined to specific industries like textiles (Mokyr, 2008). 1800 predates the large scale adoption of steam engines and the invention of railways. The industries that accounted for growth over the second half of the 19th century—especially steel, secondary metals industries like steel shipbuilding, and chemicals—did not yet meaningfully exist.

The logic of examining surnames is that they trace out historic patterns of migration. For instance Porteous (1982) examines the surname "Mel"—which likely derives from the Danish word for meal—and finds that it has been concentrated since at least the 16th century in the North East of England, which experienced large scale Norse settlement in the early middle ages. Kandt, Cheshire and Longley (2016) find that geographical clusters of surnames in the UK correspond to geographical clusters estimated using genetic data.

Given a matrix in which each row is a vector of surname shares for a given district, we

this period). To facilitate the use of county-level data and fixed effects, we split registration districts falling into multiple counties, so that each unit falls into one county. We do so by allocating parishes to registration district-county units.

use the spectral clustering algorithm developed by John et al. (2020) to cluster districts. This algorithm first calculates a kernel similarity matrix between districts and then runs a Gaussian Mixture Model on the eigenvectors of the kernel matrix. It also uses these eigenvectors to estimate the optimal number of clusters. The resulting clusters, depicted in Figure 3, correspond to historical regions. For instance the W region of Figure 3 corresponds to the distinctive "West Country," and combined with the S E region traces out the boundaries of the Anglo Saxon Kingdom of Wessex, while the E region corresponds to East Anglia.

These regional boundaries were also predictive of behavior. The number migrating in 1851 between districts allocated to different clusters was 15% lower, even when comparing pairs of districts nested within the same pair of counties (Table B.1).[5] Comparing pairs of individuals resident in the same parish, those born in different clusters were less likely to be married (Table B.2), even when we control for the higher rate of marriage among those born in the same district. These validation tests increase our confidence that the estimated clusters correspond to meaningful cultural groupings. Clearly, there exist alternative ways to divide up England and Wales according to culture. In Appendix H.6, we verify that our results are robust to alternative cultural clusters based on linguistic profiles and administrative geographies.[6]

3.3 *Measuring Cultural Choices*

We proxy for cultural choices using first names. We allocate names to cultural groups using data on the earlier generation in our data, those born between 1841 and 1860, who were aged 51–70 in 1911. We can then examine the types of names—in terms of their local identity content—that were given to the younger generation, which is the focus of our main analysis in Sections 5-7.

Using the frequencies of names among the 1841-1860 cohort in different registration districts, and an allocation of registration districts to a set of cultural clusters, we calculate culture name scores for each name i and each cluster k:

$$\textit{Culture Name Score}_{i,k} = \frac{P(\text{name} = i|\text{culture} = k)}{P(\text{name} = i|\text{culture} = k) + P(\text{name} = i|\text{culture} \neq k)}$$

where culture in the above expression refers to whether an individual was born in a given

[5] Day (2023) uncovers a similar set of cultural regions by applying a community detection algorithm on migration flows between 1851 and 1911. Migration is a choice outcome in our model, which is why we prefer to rely on pre-determined cultural markers like surnames to identify cultural clusters prior to 1851.

[6] We consider both of these alternative approaches to be inferior to reliance on surnames. Linguistic profiles come from written records and do not capture the granularity of spoken local dialects, while boundaries of administrative units are not necessarily drawn with cultural differences in mind.

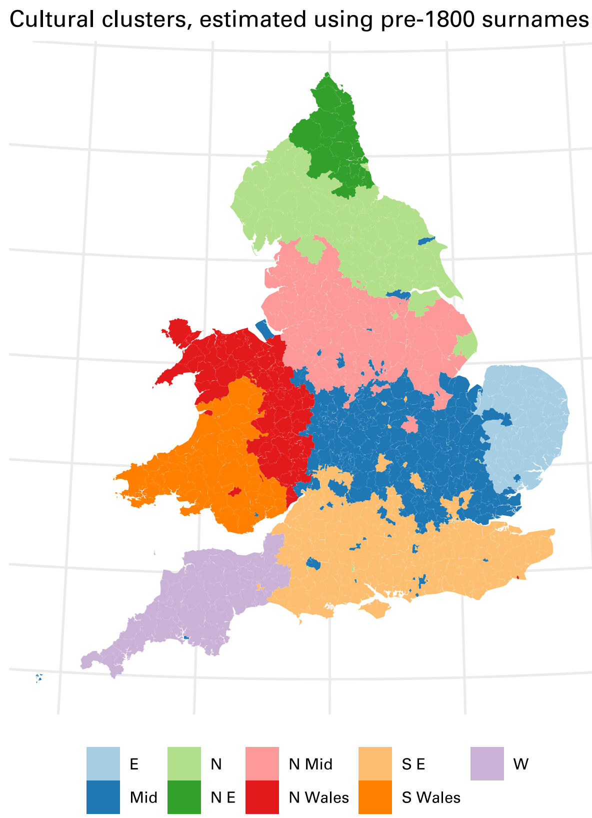

Figure 3: Map of cultural clusters based on surname frequencies

Clusters from running spectral clustering algorithm on surname frequencies of individuals present in the 1851 census who were born before 1851.

cultural cluster.[7] Name scores of this variety have been used widely to study racial identity (Fryer Jr and Levitt, 2004), immigrant assimilation (Fouka, 2019), and nation-building (Bazzi et al., 2019). Table B.3 reports the highest-scoring names for each cluster. These names trace out linguistic divides—Welsh names like Gwenllian and Llewellyn score highest in the Welsh clusters. High-scoring names also reflect local histories. Cuthbert, the namesake of the patron saint of Northumbria in the northeast of England, is the highest scoring name for the Northeast English cluster. Sutcliffe, a surname which originates from a number of villages in Yorkshire, is the highest scoring name for the North-Mid English cluster.

We take names to be a summary indicator of other changes in markers and behaviors that could be reflective of identity change, but are harder to measure in our context. Research indicates that names are a good proxy of latent attributes such as values and group attachments and strongly correlate with behavioral measures of identity (Fouka, 2020; Bazzi, Fiszbein and Gebresilasse, 2020). To the extent that we can validate it, this connection also holds true

[7]We standardize name spellings using Metaphone codes.

in our data. Table B.4 shows that those with higher name scores for the two Welsh clusters were more likely to speak Welsh and less likely to speak English. Name choices by parents can then reasonably be inferred to correlate with other meaningful acculturation decisions, such as language investments and therefore represent an informative proxy of overall identity change.

With culture name scores at hand, we can quantify identity change by examining names given to subsequent generations and the cultural clusters those names were associated with. This allows us to trace the popularity of different local cultures over time.

3.4 *Patterns of identity change*

The economic transformations experienced by England and Wales between 1851 and 1911 were accompanied by equally dramatic changes to the cultural map. We document three stylized facts about the process of cultural change during this period.

Fact 1. The culture of the Southeast of England increased in popularity.

The first and most notable development is cultural homogenization towards the culture of the Southeast. Figure 4 maps the cluster with the highest average name score among those born in each registration district in the periods 1851–1860 and 1901–1910. The culture of the Southeastern cluster increased in popularity relative to other clusters; entire cultures disappeared from the map by the 1900s. Cultural homogenization, as measured with naming, in the late 19th century coincides with other forms of observed cultural homogenization. An overview of linguistic change in 19th century England notes that regional dialects declined in prevalence from the 1870s onwards, and were replaced by a linguistic standard based on the dialect of the area around London (Görlach, 1999). The apparent decline of dialects motivated intellectuals, like the novelist Thomas Hardy, to begin cataloging regional dialects. They linked the death of dialect to migration out of rural regions.

Fact 2. Cultural change followed a center-periphery gradient.

The second noteworthy pattern is that identity change was stronger in the periphery. While the popularity of the Southeastern culture increased everywhere, this increase was larger in regions further away from London. This can be seen in the left panel of Figure 5, which is a binned scatterplot of the relationship between the growth in popularity of the Southeastern culture between those born 1851–1860 and those born 1901–1910 (y-axis) and distance from London (x-axis). Both the intercept and slope are positive capturing overall convergence to the Southeast and a center-periphery gradient.The mirror image of this pattern is the decline of local cultures, which was also steeper in peripheral regions (Figure A.4).

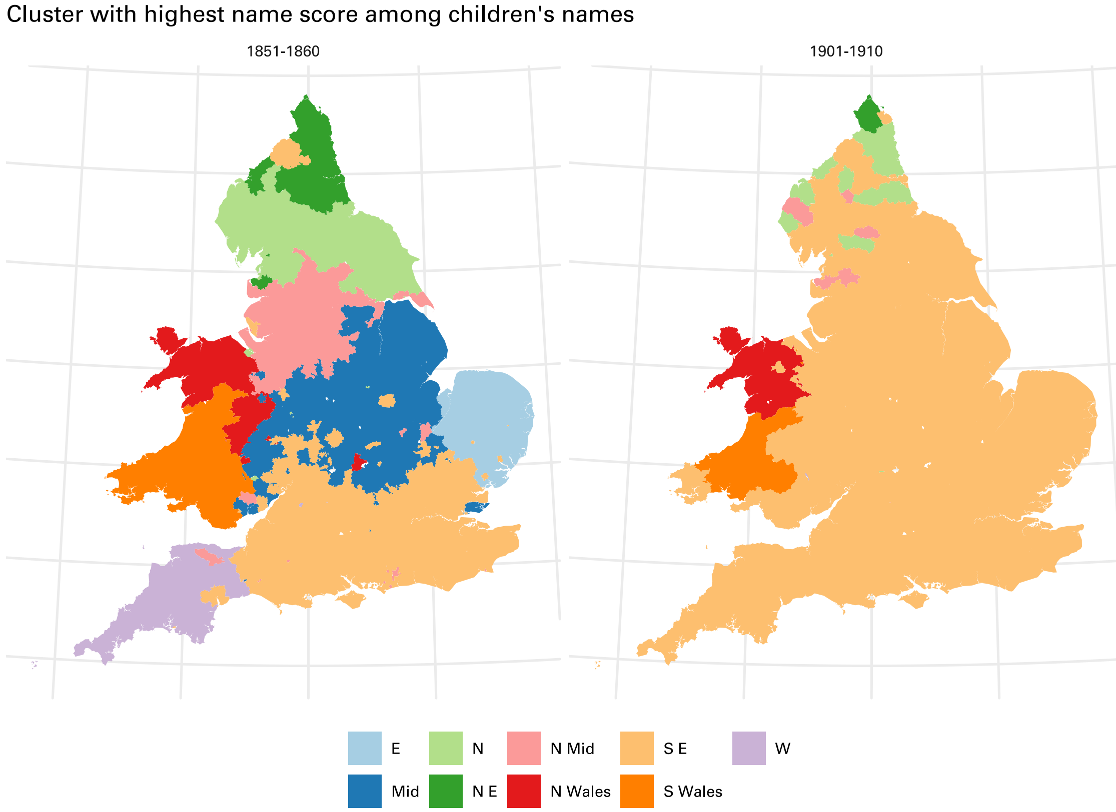

Figure 4: Changing spatial distribution of culture-specific names

These maps show the cluster with the highest name score among children born in each district, 1851–1860 (left) and 1901–1910 (right). We calculate the average name score for each cluster for those born in the district, and then record the cluster with the highest average score.

Fact 3. Peripheries with coal resisted the overall tendency of peripheries to lose their local cultures.

The decline of local cultures in peripheral locations was heterogeneous. The right-hand panel of Figure 5 plots the change in the log share given names most associated with the home culture against distance to London. Local cultures declined more in districts further from London, but this change was moderated by the presence of coal, which proxies for the development of local industry. While more peripheral districts tended to lose their culture at a higher rate, this is not true for districts that contained coal deposits. For those districts, the relationship between average distance and loss of home culture is flatter. Conversely, while the presence of coal is associated with a decline in local cultures, this appears to be mainly driven by districts near the center of the country. Peripheral districts with coal remained relative cultural holdouts.

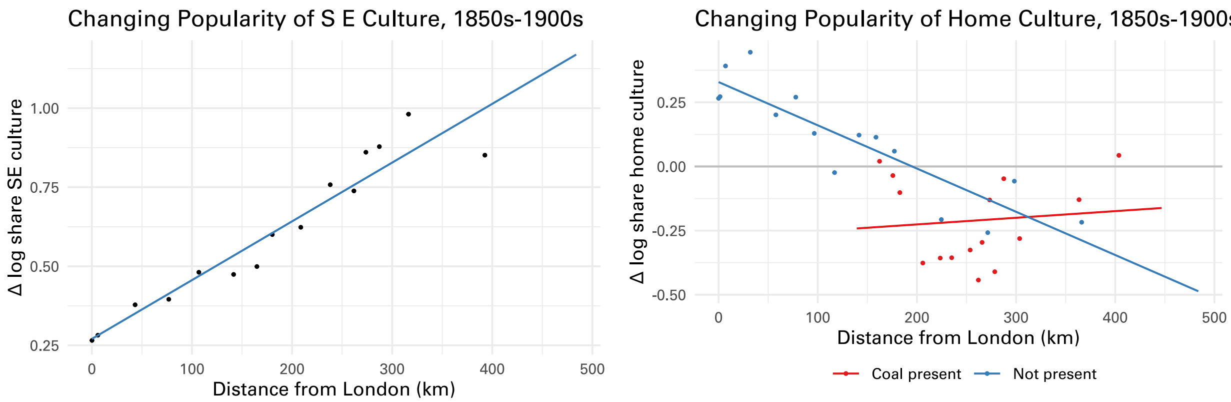

Figure 5: Distance to London, coal, and the changing popularity of Southeast English and home cultures over the late 19th century

The figure on the left shows the change in the log share first names most associated with the Southeast culture, comparing those born between 1851–1860 and 1901–1910, against the distance from the district centroid to the City of London. The right panel shows the change in the log share first names most associated with the home culture over the same period, subset by district containing coal deposits, plotted against distance to London. In both figures, observations are weighted by the number allocated names in the 1851–1860 cohort. Table B.5 presents the underlying regression estimates.

4 MIGRATION AS A LINK BETWEEN ECONOMIC AND CULTURAL CHANGE

Migration was a likely driver of the patterns of identity change we document. During the period we examine, as a result of the technological and other economic developments described earlier, the volume of migration flows increased drastically. Figure 6 shows the increase in lifetime migration out of districts and clusters of birth between 1851 and 1911. The fraction migrating from their cluster of birth increased from 18% in 1851 to 24% in 1911. Migrants moved towards industrializing centers. Models (1) and (2) in Table 1 regress changes in (log) rates of in- and out-migration between the 1850s and 1900s against changes in (log) manufacturing workers. Locations that industrialized experienced more in-migration and less out-migration.

These changes in migration rates, in turn, are associated with cultural change. Models (3) and (4) of Table 1 show that increases in the log share of both in- and out-migrants are correlated with increases in the log share given names most associated with the Southeast cluster. The pattern is similar, but reversed, for home cultures (Appendix Table B.6); both in- and out-migration were associated with declines in local identities. In-migration likely brought in people with different cultures, increasing the prevalence of other cultures like the Southeast and reducing that of the home culture. At the same time, out-migration may have created stronger incentives to abandon local identities and assimilate into the cultures of likely destinations. These patterns also imply an ambiguous relationship between industrialization

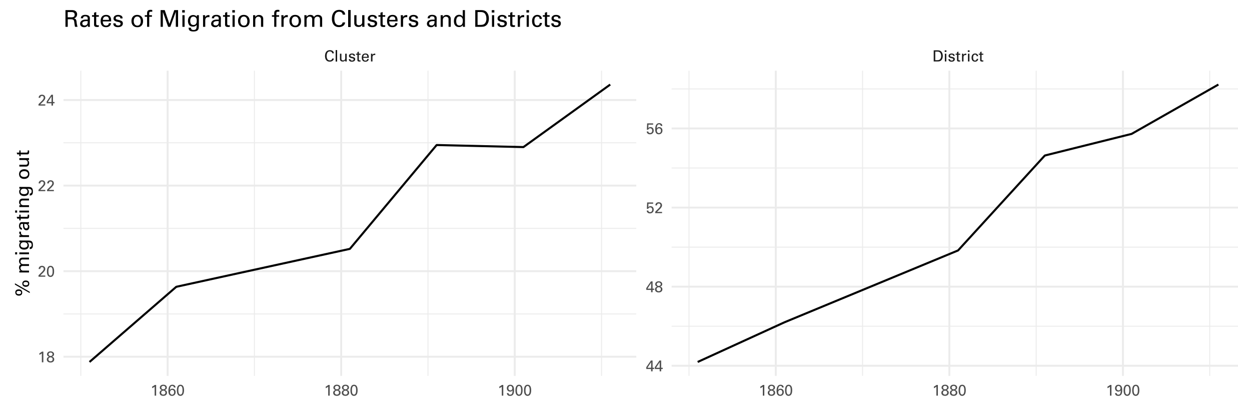

Figure 6: Increasing rates of migration 1851–1911

The figure plots the share of people living outside their cluster or district of birth at the time of the census, in the 1851–1911 censuses.

and identity change. Since industrialization both encourages in-migration and discourages out-migration (columns (1) and (2) of Table 1) it creates two offsetting effects, one pushing towards adoption of the Southeast culture and one towards preservation of local cultures.[8]

Migration and cultural choices also influenced one another at the individual level. Table 2 shows that people were more likely to migrate to locations that were culturally similar to them. Columns (1) and (2) display the results of regressions in which the dependent variable is the share of people with a given name born in a given district migrating to a district allocated to a given cluster, and the independent variable is their name score for that cluster. A name more associated with a particular culture is strongly predictive of migration to that cluster, even in specifications which adjust for the propensity of people born in a given district to migrate to a given cluster, and thus compare individuals to others born in the same district with less culturally-aligned names.

The migration histories of people named Owen, born in Anglesey in the north-west corner of Wales, illustrate these patterns (Figure A.3 shows the location of Anglesey). Owen receives name scores of 0.96 for North Wales and 0.79 for South Wales—which is unsurprising given it is the (anglicized) name of the medieval Welsh hero Owain Glyndŵr—but 0.29 for Southeast England. We would expect Owens to remain in Wales, and they did: 78% of Owens born in Anglesey remained in North Wales, 10% migrated to South Wales, and 2% migrated to Southeast England.

It was not just that cultural choices were predictive of migration; migration opportunities also influenced cultural choices. Columns (3)–(5) examine the relationship between the log

[8]Indeed, in Table B.7 we do not find a robust relationship between changes in manufacturing employment and the prevalence of the Southeast culture. Our model allows us to examine which effect dominates in different contexts, a question to which we return in Section 8.

	Δ in-mig.	Δ out-mig.	Δ ln share S E culture	
	(1)	(2)	(3)	(4)
Δ ln Mf workers	0.331	−0.132		
	(0.085)	(0.034)		
Δ ln in-migrants / pop			0.066	
			(0.018)	
Δ ln out-migrants / pop				0.128
				(0.050)
N	815	816	809	810
R^2	0.065	0.060	0.062	0.029

This table presents the results of district-level regressions linking cultural change, migration and industrialization. In model (1), the dependent variable is the change in the log number of people over 16 born outside the cluster living in the district, divided by the number born in the district, between 1851 and 1901, in (2) the change in the log number over 16 born inside the district living outside the cluster, divided by the number born in the district, in (3)–(4), the dependent variable is the change in the log share assigned names most associated with the Southeast cluster, between the cohort born 1851–1860 and that born 1901–1910. In (1) and (2), the independent variable is the change in the log number of manufacturing workers, between 1851 and 1901. The independent variable in (3) is the dependent variable in (1). In (4), the dependent variable in (2). In (1)–(2) observations are weighted by 1851 population, in (3)–(4), by the number allocated namescores in the 1851–1860 cohort. Robust standard errors in parentheses.

Table 1: Relationship between migration, industrialization, and the rise of the Southeast culture

share migrating from a given district to districts allocated to particular clusters and the log share in that district given names most associated with that culture. Both variables are measured using those born in a given district between 1861 and 1895. Returning to Anglesey, 75% of those born in Anglesey remained in North Wales, while 8% migrated to South Wales, and 1% migrated to Southeast England. Angelesey's naming patterns reflect this orientation: names most associated with North Wales and South Wales each account for 20% of the total, names associated with Southeast England account for 7%.

A natural concern, especially given the results in columns (1) and (2), is simultaneity. (5) therefore instruments for district-by-culture migration flows using migration flows to different regions predicted simply based on the presence of coal deposits in the destinations and geographic distance. We describe the construction of this instrument in more detail in Appendix D.1. To correct for the fact that districts closer to districts associated with a particular cluster will be predicted more migration to that cluster due to proximity, not the pull of coal, we follow Borusyak and Hull (2020) and permute the vector of coal allocations 1,000 times, and on each permutation calculate predicted log migration shares to each cluster from each district. We then subtract the average of these permuted instruments from our instrument, and add district and cluster fixed effects. In both the OLS and TSLS specifications,

	ln share migrating			ln share names	
	(1)	(2)	(3)	(4)	(5)
name score	2.323	0.716			
	(0.122)	(0.034)			
recentered coal-predicted					
ln share migrating			2.586		
			(0.419)		
ln share migrating				0.122	0.147
				(0.003)	(0.021)
Name x district FE	x	x			
District FE			x	x	x
Cluster FE	x		x	x	x
Cluster x district FE		x			
Model	OLS	OLS	First Stage	OLS	TSLS
First stage F-stat					38.1
N	840394	840394	7272	7272	7272
R^2	0.521	0.934	0.571	0.965	0.964

This table presents evidence of the relationship between cultural naming choices and migration. Models (1) and (2) are estimated at the name-district-cultural cluster level: the dependent variable is the log share of people with a given name born in a given district migrating to a district in a given cultural cluster. The independent variable is the name score for that name for the destination cultural cluster. Both models include fixed effects for the name-district of birth combination, (1) includes fixed effects for the destination cluster, (2) interacts these with the district of birth. (1) and (2) are weighted by the number of people with each name born in each district. Models (3)–(5) are estimated at the district-cultural cluster level. In (4) and (5) the dependent variable is the log share given names most associated with the cultural cluster, and the independent variable is the log share of individuals migrating from the district to that cluster. In (5) this is instrumented for with the log share of migrants predicted by the location of coal deposits in a gravity model, recentered following Borusyak and Hull (2023). We permute the vector of coal deposits across district, calculate predicted log share of migrants under each permutation, and substract the mean of this from the instrument. (3) shows the first stage. (3)–(5) all include district and cluster fixed effects, and are weighted by the number of individuals with name scores born in each district. Standard errors clustered by district in parentheses.

Table 2: Relationship between migration and naming patterns

we find a strong positive relationship between migrants from a location to a cluster and the prevalence of names associated with that cluster. A 1% increase in migrants to a cluster is associated with around a 0.15% increase in the share given names most associated with the cluster. This suggests that parents made naming decisions in anticipation of the likely migration destinations for their children. Taken together, the results of Table 2 suggest that migration flows and observed identities should be viewed as outcomes of decisions that were jointly determined.

5 A MODEL OF MIGRATION AND IDENTITY CHOICE

Building on the results of Table 2, we model migration and identity decisions explicitly in a spatial equilibrium model. Two considerations motivate the use of a formal model and structural estimation—as opposed to reduced form regressions—to make sense of patterns of cultural change via migration. First, the model provides a principled way to decompose the quantitative contribution of two core developments of Britain's modernization—changing patterns of economic activity and falling migration costs—to the patterns of cultural change that we observe. Second, reduced form approaches are methodologically inappropriate for the setting we analyze. One reason is interdependence across regions. How culture will develop in one district depends on what happens everywhere else. Imposing theoretical structure allows us to estimate spillovers, rather than viewing them as a threat to estimation, as would be the case in linear regression models. Additionally, in a reduced form analysis, changes such as drops in migration costs, which affected all districts over the period of study, can only be studied in terms of their differential effects across locations, but not their overall level. Structural estimation allows us to quantify the effect of such global changes.

In the model, agents, who are distinguished by their location of origin and culture, decide where to live. They choose locations in part due to economic incentives, and in part due to a preference to reside with others of the same culture. While we anticipate, and indeed find, that this preference is positive, our model is relatively general and treats the direction and magnitude of that preference as quantities to be estimated. Agents also have culture- and individual-level idiosyncratic preferences for specific locations. Their parents choose to raise them in a particular culture, not knowing fully what these individual-level preferences will be. We model the decision of parents as a choice over the assignment of an identity marker—corresponding to name choice in our data—which, however, we take to proxy for additional cultural investments the parents may make in a specific identity, such as in language or education. Parents care about the expected welfare of their child under different cultures, but also have their own preferences for specific cultures.

5.1 *Setup*

There are N regions, and K cultures. We will use subscripts d (for destination), o (for origin), and j to index regions, and superscripts k and l to index cultures. There are two sets of agents: parents and children. Parents choose which culture to raise their offspring in; the younger generation then, having received an assigned identity from their parents, decides where to live and work. When choosing where to migrate, the utility person i of culture k

born in location o receives from migrating to destination d is

$$u^k_{odi} \equiv v_d \delta_{od} \left(m^k_d\right)^{\frac{\alpha}{\theta}} \xi^k_d \varepsilon_{di}.$$

v_d is the real wage in d, an attribute of the location that is experienced equally by members of all cultures.[9] δ_{od} is an inverse cost of migrating from o to d, and captures the idea that it is easier to migrate from a given origin to some destinations than others. m^k_d is the number of people who end up in d drawn from culture k, and α is the "homophily elasticity," the elasticity of migration choices to the population of the same culture. If α is positive, people have a preference for locating in places with more people of the same culture as them. ξ^k_d is a taste parameter, known to the agents though unobserved by the econometrician, for location d for all members of culture k. This "cultural amenity" variable represents unmodeled preferences for certain places by members of certain cultures, distinct from the benefit of being with other members of the same culture. One such preference would be a desire to live in an ancestral homeland. ε_{di} is a preference shock drawn iid from a Fréchet distribution with shape θ over individual-by-destination pairs. This preference shock accounts for individual heterogeneity in migration decisions.[10]

Each agent migrates to the place that gives her the greatest utility, inclusive of her preference shock. It follows from the properties of the Fréchet distribution (McFadden, 1974; Hsieh et al., 2019; Bryan and Morten, 2019), that the number of members of culture k in o migrating to d is given by

$$m^k_{od} = \frac{\left(v_d \delta_{od} \xi^k_d\right)^{\theta} \left(m^k_d\right)^{\alpha}}{\sum_{j=1}^{N} \left(v_j \delta_{oj} \xi^k_j\right)^{\theta} \left(m^k_j\right)^{\alpha}} n^k_o, \tag{1}$$

where n^k_o is the number of people originating in o with culture k. This expression implies that more people will migrate to a location if wages are high (v_d is large), if it is near to places with large numbers of potential migrants ($\delta^{\theta}_{od} n^k_o$ is large), if it is more popular specifically with the cultures of these migrants ($\left(\xi^k_d\right)^{\theta} \left(m^k_d\right)^{\alpha}$ is large), and if there are few other compelling locations competing for the same migrants (the denominator is small).

It also follows from the properties of the Fréchet distribution that expected utility—prior

[9]While we do not explicitly model the economy, one can think of v_d as capturing factors like productivity and resource endowments that influence output per worker. This functional form could emerge if in each location a homogeneous and freely-traded numeraire good is produced using labor under perfect competition and constant returns to scale, with productivity v_d.

[10]The θ parameter governs the responsiveness of migration decisions to the varying utilities of different destinations. Our empirical strategy does not separately identify θ, but we do not need to do so to estimate counterfactuals.

to the realization of the ε_{di} shock—for members of culture k in o, is

$$\left(\sum_{j=1}^{N}\left(v_j \delta_{oj} \xi_j^k\right)^{\theta}\left(m_j^k\right)^{\alpha}\right)^{\frac{1}{\theta}} \Gamma\left(\frac{\theta-1}{\theta}\right),$$

where $\Gamma(\cdot)$ is the Gamma function. Ignoring the Γ constant, we define the variable Ω_o^k:

$$\Omega_o^k \equiv \sum_{j=1}^{N}\left(v_j \delta_{oj} \xi_j^k\right)^{\theta}\left(m_j^k\right)^{\alpha}. \tag{2}$$

An agent's parents choose to assign her to one of K cultures not knowing her vector of ε_{di} shocks. They make this cultural decision trading off the benefits of migration and cultural homogeneity against a place-specific attachment to different cultures. This assumption is similar to, but simpler than, the "imperfect empathy" of Bisin and Verdier (2001). Instead of parents evaluating their children's choices with their own utility function, they simply have preferences over their children's cultural identities that depend partly on the children's welfare and partly on the parent's tastes. More formally, the utility the parent of agent i in o receives from choosing culture k is

$$\bar{u}_{oi}^k \equiv \left(\Omega_o^k\right)^{\frac{1}{\theta}}\left(\psi_o^k\right)^{\frac{1}{\varphi\theta}} \iota_i^k$$

where ψ_o^k is the "cultural transmission taste," a place-specific bias that parents have for assigning their children a given culture, and ι_i^k is an individual-by-culture Fréchet shock, with shape $\varphi\theta$. It follows again from the properties of the Fréchet distribution that the share choosing culture k in o is

$$\sigma_o^k = \frac{\left(\Omega_o^k\right)^{\varphi}\psi_o^k}{\sum_{l=1}^{K}\left(\Omega_o^l\right)^{\varphi}\psi_o^l}. \tag{3}$$

Here φ is the "culture elasticity," the parameter that determines how responsive cultural choices are to migration opportunities. A larger value of φ implies that the distribution of ι_i^k is less dispersed, meaning that differences in individual ι_i^k shocks for different cultures will be small in relation to Ω_o^k, and consequently choices will largely depend upon Ω_o^k, the expected utility experienced by the child raised in that culture. As φ tends towards infinity, we would expect all parents in a given location to choose the same culture for their children.

5.2 *Equilibrium*

An equilibrium in our model is a vector of culture-by-origin choices and culture-by-origin-by-destination migration flows such that equations (1) and (3) are satisfied. This implies the

Table 3: Structural Parameters

	Reference	Interpretation	Level	Estimation
α	Homophily elasticity	Responsiveness of migration decisions to other members of the same culture	Universal	Two-stage least squares
φ	Culture elasticity	Responsiveness of cultural choices to migration opportunities	Universal	Instrumental variables Poisson regression
δ_{od}^{θ}	Inverse migration cost	Inverse cost of migrating from o to d	Origin-by-destination	Gravity regression of migration flows against distance and fixed effects
v_d^{θ}	Real wage	Non-cultural incentive to migrate to location d	Destination	Backed out from fixed effects in gravity regression
$\left(\xi_d^k\right)^{\theta}$	Cultural amenity	Culture k-specific taste for migrating to location d	Culture-by-destination	Backed out from fixed effects in gravity regression
ψ_o^k	Cultural transmission taste	Taste of parents in o for assigning children culture k	Culture-by-origin	Backed out from cultural choices in origin locations
Ω_o^k	Expected utility	Expected value of belonging to culture k if born in location o	Culture-by-origin	Function of $\alpha, v_d, \delta_{od}, \xi_d^k$ and migration flows

assumption of optimal behavior for each parent, fixing the actions of all other parents.

6 TAKING THE MODEL TO THE DATA

In the model, migration flows and cultural choices are generated by agents optimizing according to preferences defined by the structural parameters $\alpha, \varphi, v_d^{\theta}, \delta_{od}^{\theta}, \left(\xi_d^k\right)^{\theta}$, and ψ_o^k, for $o, d \in \{1, ..., N\}$, and $k \in \{1, ..., K\}$. We treat the observed migration flows and cultural choices as an equilibrium realization of the model, and back out the structural parameters in three stages. Table 3 provides an overview of the different parameters and the estimation steps we follow to recover them from our data.

6.1 *Estimating Migration Pull Factors and Costs*

In the first stage, we estimate the bundle of features that make a given location attract migrants from a given culture. Taking logarithms of (1) gives

$$\underbrace{\ln m_{od}^k}_{\ln \text{migrants from } o \text{ to } d \text{ of culture } k} = \underbrace{\ln\left(v_d^\theta \left(\xi_d^k\right)^\theta \left(m_d^k\right)^\alpha\right)}_{\text{Destination-culture FE}} + \underbrace{\theta \ln \delta_{od}}_{\ln \text{distance}} + \underbrace{\ln\left(n_o^k/\Omega_o^k\right)}_{\text{Origin-culture FE}} .$$

This equation implies that in a regression of log migration flows against some measure of migration costs and origin-by-culture and destination-by-culture fixed effects, the destination-by-culture fixed effect recovers the (log) expected utility of moving to a given location for members of each culture. We parameterize migration costs as a log-linear function of geographic distance and an indicator that the origin equals the destination, $\delta_{od}^\theta = \text{distance}_{od}^{\beta_1} \exp\left(\beta_2 \mathbf{1}_{\{o=d\}}\right)$. This choice follows scholarship on the gravity structure of migration, which indicates that migration flows are strongly related to distance. Inserting this parameterization of δ_{od}^θ into the equation above gives an estimating equation:

$$\ln m_{od}^k = \gamma_d^k + \beta_1 \ln \text{distance}_{od} + \beta_2 \mathbf{1}_{\{o=d\}} + \gamma_o^k + \varepsilon_{od}^k.$$

We estimate this model by Poisson Pseudo-Maximum Likelihood (Silva and Tenreyro, 2006; Fally, 2015). The error term ε_{od}^k here represents unobserved shocks to origin-by-destination migration costs. We assume these are transitory and do not affect choices of culture. The destination-by-culture fixed effect γ_d^k estimates $\ln\left(v_d^\theta \left(m_d^k\right)^\alpha \left(\xi_d^k\right)^\theta\right)$.

6.2 *Disaggregating Migration Pull Factors*

This estimated bundle of features that attract migrants to a location includes both exogenous cultural amenities ξ_d^k, and endogenous factors related to the presence of other members of a given culture in that location, $\left(m_d^k\right)^\alpha$, as well as the real wage that attracts members of all cultures to a location, v_d^θ. In the second stage we disentangle these three components. Writing the definition of γ_d^k in separate terms gives a regression equation:

$$\underbrace{\gamma_d^k}_{\text{Destination-culture FE}} = \alpha \underbrace{\ln m_d^k}_{\ln \text{destination-culture population}} + \underbrace{\ln v_d^\theta}_{\text{Destination FE}} + \underbrace{\ln\left(\xi_d^k\right)^\theta}_{\text{Error}} . \tag{4}$$

In a regression of the estimated destination-by-culture fixed effects against log post-migration population of a given culture in a given destination and a destination fixed effect, the coefficient on log migration flows recovers α, the homophily elasticity, the destination fixed effect recovers (log) v_d^θ, the real wage, and the regression residual recovers the log of $\left(\xi_d^k\right)^\theta$, the cultural amenity.

Estimating (4) is complicated by migration flows being endogenous to cultural amenities. If a feature of place d makes it more attractive to members of a culture k (increasing ξ_d^k), more members of k will migrate to d, increasing m_d^k. Thus the independent variable $\ln m_d^k$ will be positively correlated with the error term $\theta \ln \xi_d^k$. An additional concern is that cultural transmission tastes, ψ_o^k, that influence whether parents in a given place raise their children in a given culture, might be correlated with cultural amenities, ξ_d^k, that influence whether members of a culture choose to migrate there. As migration flows decline with distance, any factor that increases the number of people of a culture originating in a location will also increase the number migrating to that location.

We need an instrument for culture-destination migration flows that is uncorrelated with cultural amenities. We develop such an instrument building on our analysis in Table 2 linking the distribution of coal deposits across regions to cultural choices in the origin locations. We use distance and coal deposits to predict migration flows to destinations and thus predict how shocks to cultural choices in origin locations propagate through to destination populations. Finally, we recenter the instrument following Borusyak and Hull (2020). We discuss the construction of this instrument in Appendix D.3. The key assumption for identification is that this instrument is uncorrelated with factors distinct from the destination population that attract members of a culture to a given location. While we cannot directly verify that assumption, we can verify that the instrument is uncorrelated with observable measures of cultural similarity between locations. We discuss this check in Appendix D.4.

We estimate α by two-stage least squares with fixed effects for destinations and for the interaction of the destination culture (based on the surname-based cluster it is allocated to) and the culture in question. This latter set of fixed effects account for the concern that the surname cluster a district is in might be correlated with cultural amenities.

Given an estimate of α, we can then regress $\gamma_d^k - \alpha \ln m_d^k$ against destination fixed effects to recover v_d^θ (the exponential of the destination fixed effect) and $\left(\xi_d^k\right)^\theta$ (the exponential of the residual).

6.3 *Estimating Cultural Choice Parameters*

In the third stage we estimate the parameters linking cultural choices to migration opportunities. Taking estimates of inverse migration costs, δ_{od}^θ, and the bundle of factors that attract people to a given location, $\left(v_d \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha$, from 6.1, we can calculate Ω_o^k, the expected utility in culture k of one born in o, according to equation (2).

Taking logarithms of equation (3) gives the following regression equation:

$$\underbrace{\ln \sigma_o^k}_{\ln \text{share assigned } k \text{ in } o} = \varphi \ln \Omega_o^k - \underbrace{\ln\left(\sum_{l=1}^{K} \left(\Omega_o^l\right)^{\varphi} \psi_o^l\right)}_{\text{place FE}} + \underbrace{\ln \psi_o^k}_{\text{Error}}. \tag{5}$$

This equation implies that in a regression of the log of the share choosing culture k in location o (σ_o^k) against the log of Ω_o^k and a location fixed effect, the coefficient on $\ln \Omega_o^k$ corresponds to φ, and the residual corresponds to the logarithm of ψ_o^k.

Estimating (5) is complicated by endogeneity. An exogenous preference for choosing a given culture in a given location—ψ_o^k, the error term in equation (5)—causes there to be more people of that culture resident in proximate locations. As Ω_o^k is increasing in the number of members of culture k migrating near o, a local increase in the propensity to choose culture k should increase Ω_o^k, the independent variable. A secondary source of endogeneity is that cultural transmission tastes ψ_o^k may be correlated with cultural amenities ξ_o^k, which also features in the independent variable Ω_o^k.

As in the previous section, we use the interaction between coal deposits encouraging migration, and the spatial distribution of historical cultures, to develop an instrument. We predict the migration attraction of each location using coal deposits, and calculate the migration pull of each culture in each origin location by allocating destination locations to cultures based on our clustering of pre-1800 surnames. Appendix D.2 provides further detail. We recenter the instrument following Borusyak and Hull (2020). Appendix D.4 confirms that this instrument is also uncorrelated with observed cultural similarity. We use Poisson Pseudo-Maximum Likelihood to estimate φ because it can account for zeros in the cultural choice share (σ_o) and fits the multinomial logit structure of the cultural choice (Baker, 1994).[11] As recommended by Wooldridge (2010), we use the control function approach for instrumental variables estimation. We first run an OLS regression of $\ln \Omega_o^k$ against the instrument and fixed effects, and include the residuals from this first stage regression in the second stage. We include fixed effects for origins and for the combinations of the surname culture each location is allocated to and the culture in question, as above. Given an estimate of φ, we can back out ψ_o^k.

[11] To see how Equation (3) corresponds to multinomial logit, note that it can be rewritten as $\sigma_o^k = \frac{\exp(\varphi \ln \Omega_o^k + \ln \psi_o^k)}{\sum_{l=1}^{K} \exp(\varphi \ln \Omega_o^l + \ln \psi_o^l)}$

6.4 *Estimation*

Our estimation routine proceeds in multiple stages. First, we allocate individuals to cultures using first names. For each name, we compute its score for each cultural cluster using the process described in Section 3.3 and then allocate a name to a particular culture if its name score is highest for that culture. For instance, we consider the name David—the patron saint of Wales—to reflect the South Wales culture, because its name score for that cluster is 0.92, while its two next highest name scores are 0.67 for the North Wales cluster and 0.44 for the Eastern England cluster.

Given name scores calculated from those born between 1841 and 1860 and an allocation of names to cultural clusters we can then measure origin-by-culture-by-destination migration flows for the generation born between 1861 and 1895, that of working-age adults in 1911. Specifically, we measure migration flows by computing the number of individuals from a given culture (classified by their first name as described above) and district of birth who are observed to reside in a particular destination district. We use these migration flows to estimate destination-by-culture fixed effects and migration costs, calculate Ω_o^k, and construct the instruments necessary to estimate α and φ. Tables B.11 and B.12 provide an overview of the data and steps involved in our estimation.

6.5 *Estimates*

Table 4 shows the results of these estimation routines for α and φ. The instrumental variables estimates imply $\alpha = 0.52$ and $\varphi = 1.80$.[12] Given that the bias from endogeneity on the non-instrumented coefficients should be positive—in both cases the regressor should be positively correlated with the error term—it is encouraging that the instrumented coefficients (models 3 and 6) are smaller than the non-instrumented coefficients (2 and 5).[13]

[12] Results from Allen, Arkolakis and Li (2020) establish that a sufficient—though not necessary—condition for equilibrium uniqueness in our model is $\frac{\alpha}{1-\alpha}\max(2\varphi - 1, 1) < 1$. We present the derivation of that result in Appendix C. That condition is not satisfied for these estimated elasticities. In practice, we find that regardless of starting values—including starting values that place almost all weight on one culture—the model converges to the same equilibrium with these elasticities.

[13] We report conventional robust standard errors clustered by district in models (1)–(5) of Table 4 and cluster-bootstrap (6). In doing so we follow other papers that estimate quantitative spatial models in multiple stages (see for instance Donaldson and Hornbeck 2016; Donaldson 2018; Bryan and Morten 2019; Fajgelbaum and Redding 2022; Morten and Oliveira 2023 and Allen and Donaldson 2020). A limitation of conventional standard errors is that they do not take into account uncertainty from previous stages of the estimation. In Appendix E Table E.1 we report confidence intervals from bootstrapping the entire estimation procedure for the model elasticities. We note however that inference in quantitative spatial models is not straightforward because of dependence between all observations; to the best of our knowledge there is no scholarship on how to correctly account for such dependence.

	$\ln m_d^k$	γ_d^k		$\ln \Omega_o^k$	$\ln \sigma_o^k$	
	(1)	(2)	(3)	(4)	(5)	(6)
Recentered coal-predicted $\ln m_d^k$	0.641					
	(0.172)					
$\ln m_d^k$ [α]		0.604	0.521			
		(0.018)	(0.051)			
Recentered coal-predicted $\ln \Omega_o^k$				0.214		
				(0.035)		
$\ln \Omega_o^k$ [φ]					2.185	1.803
					(0.103)	(0.304)
First-stage residuals						0.388
						(0.285)
Model	First stage	OLS	TSLS	First stage	PPML	Control function
First stage F-stat			13.867			38.416
N	7418	7418	7418	7452	7452	7452
R^2	0.989	0.999	0.999	0.998		

This table shows estimates of α and φ. All models are estimated at the district-by-culture level. Models (1) and (4) show the first stage, (2) and (5) the non-instrumented second stage and (3) and (6) the instrumented second stage. (1), (2), and (4) are estimated by OLS, (3) by TSLS, and (5) and (6) by Poisson Pseudo-Maximum Likelihood, with (6) including the residuals from (4) as a control function. In models (2) and (3) the dependent variable is the destination-by-culture fixed effect from an origin-by-destination-by-culture gravity model, in (5) and (6) it is the log share of people born in the district assigned to the culture. All models include fixed effects for the district and the culture interacted with the historic culture of the district. Models (1)–(3) are weighted by the number of individuals at the destination, (4)–(6) by the number at the origin. Standard errors clustered by district in parentheses—in (6) these are by bootstrapping (4) and (6) with a fractional random weight bootstrap clustered by district.

Table 4: Estimates of model elasticities

6.6 *Validating the Model's Estimates of Taste and Economic Parameters*

The model produces estimates of the attractiveness to migrants of each location (v_d^θ), and of factors that made locations attractive to migrants of specific cultures (ξ_d^k), and that made parents more likely to choose specific cultures in a given location (ψ_o^k). In this subsection we check whether these estimates correlate with variables not included in the model that should be associated with these economic and taste parameters. This provides a way to validate our model.

We would expect cultural amenities, ξ_d^k, and cultural persistence tastes, ψ_o^k, to be correlated with cultural proximity. Beyond having more people of the same culture, a location may be more attractive to migrants from culture k if it has resources for activities specific to that culture, like the same type of church. Similarly, people in a location may be more likely to raise their children in a particular culture if it is relatively similar to the culture of their ancestors. In Figures 7 and 8 we examine the relationship between these district-by-culture taste parameters and the average religious distance and surname similarity of the district to districts assigned to the culture in question. For religious distance, we take the share of worshippers in each denomination in each district in the 1851 Census of Religious Worship (Southall and Ell, 2022), and calculate the Euclidean distance between each district's vector of shares. For surname similarity, we use the kernel similarity metric used by our clustering algorithm applied to the surnames of those born before 1800 in the 1851 census. We take the average distance or similarity between each district and all districts allocated to a particular culture by our clustering algorithm. Both cultural transmission tastes (ψ_o^k) and cultural amenities (ξ_d^k) are strongly correlated with religious distance, though only cultural transmission tastes strongly correlate with surname similarity. Figures 7 and 8 present binned scatterplots that residualize out cluster and district fixed effects.

We would expect more economically developed locations to have higher real wages (v_d^θ). The parameter v_d^θ captures factors common across all cultures that induce migration, especially economic opportunities. During the period we study, Britain was an industrializing economy, and so the most economically developed places, where we would expect real wages to be highest, were those that had experienced structural transformation from agriculture to manufacturing. Figure 9 shows binned scatterplots of the relationship between the shares in agriculture and manufacturing, and the log real wage, $\theta \ln v_d$. There is a strong, linear, negative relationship between the share employed in agriculture and real wages, and a positive, though non-linear, relationship between the share in manufacturing and the real wage.

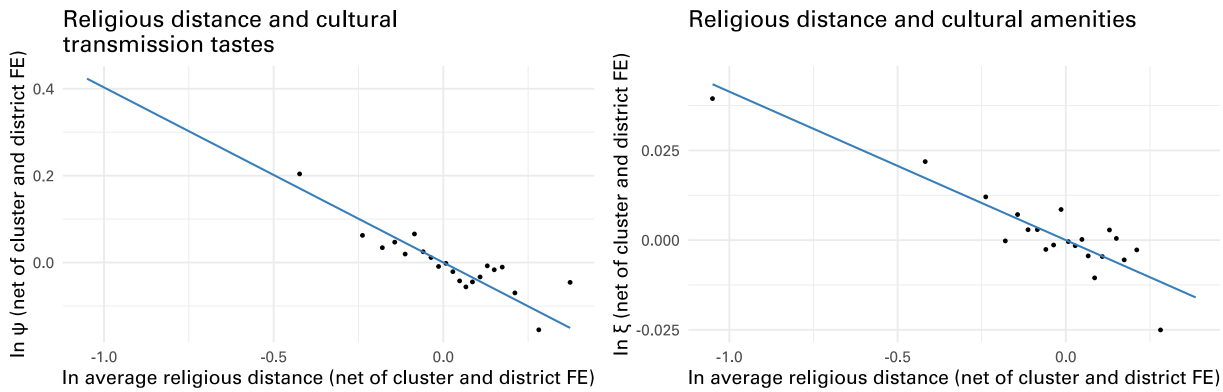

Figure 7: Binned scatterplots of cultural transmission tastes ψ_o^k, cultural amenities ξ_d^k and the average Euclidean distance of religious denomination shares in a district to those in a given cluster in 1851

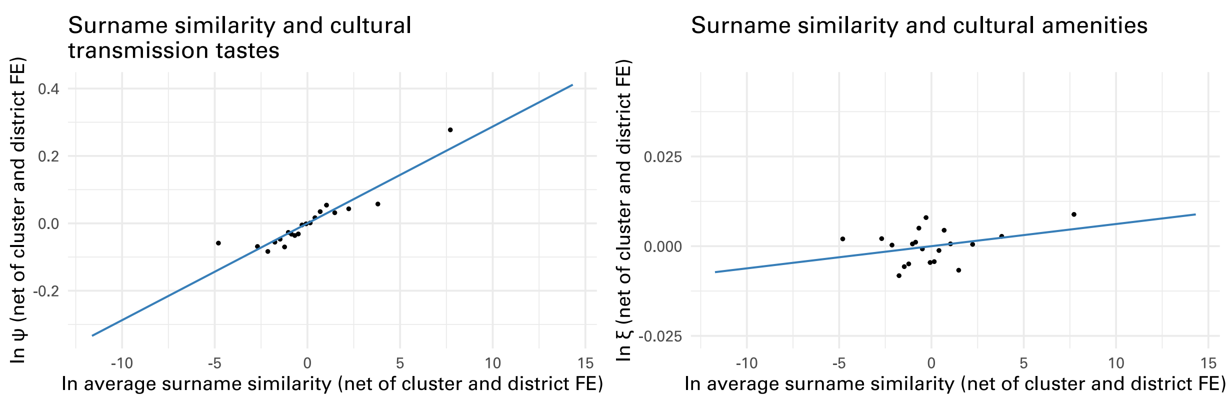

Figure 8: Binned scatterplots of cultural transmission tastes ψ, cultural amenities ξ and the average similarity of surnames in a district to those in a given cluster in 1800

7 USING THE MODEL TO UNDERSTAND PATTERNS OF CULTURAL CHANGE

Given estimates of the various economic and cultural parameters, we can solve for the cultural equilibrium under counterfactual changes to these parameters. This allows us to examine how changes to different economic fundamentals influence the cultural map. Economic fundamentals enter the model through three sets of exogenous variables, which capture facets of Britain's modernization during the Second Industrial Revolution. First, the real wage in different locations, v_d^θ, that influences migration decisions, directly corresponds to industrialization (Figure 9). Economic growth in parts of the country should increase the incentive for people to migrate there. Second, spatially-uneven economic growth also leads to spatially-uneven population growth, which alters the starting population in different locations.

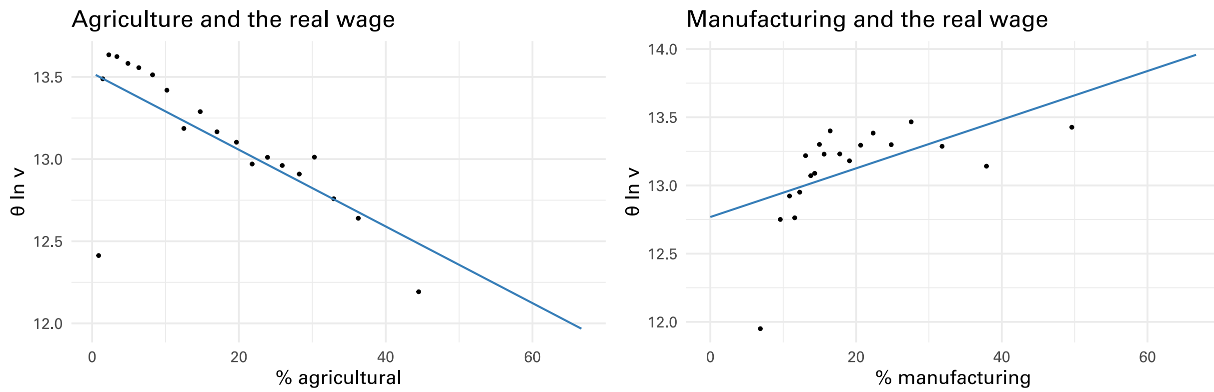

Figure 9: Binned scatterplots of shares of employment in agriculture and manufacturing and real wages $\ln v_d^{\theta}$

Third, the costs of migration change over time.

We conduct two types of counterfactual exercises. First, we use the 1851 census to calculate real wages, migration costs, and starting populations and show that the change in cultural choices predicted going from 1851 economic fundamentals to 1911 economic fundamentals matches observed changes over that period. This indicates that our model captures important mechanisms that account for observed variation in the data. Next, we study the contribution of each form of economic change separately and show that changes to real wages and migration costs account for different aspects of the patterns we documented in Section 3.

We estimate the economic components of the model using data on district-to-district lifetime migration from the 1851 Census. Our estimation routine follows that in Section 6.1, using observed migration flows to estimate pull factors and distance-related costs. The key difference relative to 6.1 is that we do not have naming data for the 1851 period and so cannot use origin-destination-culture migration flows. In practice, this means that we estimate real wages and migration costs—variables that do not vary across cultures—and need to rescale our estimates of real wages to account for the component of utility that is increasing in the number of members of a given culture. Appendix F describes the estimation in greater detail. We also use the starting populations of migrants in 1851 to provide an estimate of starting populations.

7.1 *The Model Rationalizes Patterns of Identity Change*

Our first simulation aims at examining whether the model, and the migration-related theoretical mechanism it focuses on, can recover the empirical patterns of identity change that we observe. To this end, we solve for equilibrium cultural choices given 1851 real wages, migration

	Observed cultural change ($\Delta \ln \sigma$)			
	(1)	(2)	(3)	(4)
Predicted cultural change ($\Delta \ln \sigma$)	0.195	0.173	0.182	0.115
	(0.012)	(0.012)	(0.017)	(0.016)
District FE		x		x
Cluster FE			x	x
N	7362	7362	7362	7362
R^2	0.044	0.099	0.533	0.589

This table shows OLS estimates at the district-by-cluster level. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination real wages v_d^θ, starting populations, and migration costs and the observed value for those born 1861–1895. The dependent variable is the change between the observed value for those born 1841–1860 and those born 1861–1895. Model (2) adds district fixed effects, (3) adds cluster fixed effects, (4) adds both. Standard errors clustered by district in parentheses.

Table 5: Relationship between the change in log cultural choice shares, σ, relative to the 1851 counterfactual and relative to the 1841–1860 cohort

costs, and starting populations, and 1911 cultural amenities and transmission tastes. One can think of the difference between cultural choice shares from this simulation, and 1911 observed cultural choice shares, as the model's prediction of how economic changes over the 1851–1911 period influenced cultural changes, holding fixed noneconomic factors influencing culture, which are embedded in the cultural amenity and transmission taste terms, ξ_d^k and ψ_o^k. Holding these ξ_d^k and ψ_o^k terms fixed accounts for the many other developments in this period, for instance in education, that affected cultural choices through channels other than migration.

In Table 5 we compare these predicted changes in cultural shares against observed changes in cultural shares between those born 1841–1860 and those born 1861–1895. The table reports the results of district-cluster level regressions in which the dependent variable is the change in the log share choosing each culture between the 1841–1860 and 1861–1895 generations, and the independent variable is the observed log share from 1861–1895 minus the log share predicted by our counterfactual using 1851 economic fundamentals. There is a strong positive correlation between the observed and counterfactual changes, that is robust to the addition of district and cluster fixed effects. The latter is important, as it indicates that the model successfully predicts more subtle within-cluster variation, rather than just predicting which clusters experienced increases in popularity. Figure A.5 compares the spatial distribution of the percentage change in the share choosing the Southeast English cluster between the counterfactual and observed data. While clearly not perfect, the model reproduces the spatial patterns of observed change.

What are these spatial patterns? Broadly, they correspond to the three main stylized facts we established earlier. First, the model captures the increase in the popularity of the culture of the Southeast of England (Fact 1). This can be seen in the first row of Table 6, which presents changes in the popularity of the home and Southeast cultures, and migration rates, under different counterfactual scenarios. Setting all three fundamentals to 1851 levels decreases the prevalence of the Southeast culture by around 30%.

Second, the model also reproduces the center-periphery gradient in cultural assimilation to the Southeast and home culture retention we observe in the data (Fact 2). Figures 10A and 11A plot the observed log share minus the counterfactual log share against distance from London, for the Southeast and home cultures respectively. The figures reproduce the higher increase in Southeast culture and higher loss of home culture in the peripheries.

Finally, Figure 11A recovers the heterogeneous effect of coal for home culture loss observed in Figure 5 (Fact 3). Relative to other peripheral districts, peripheries with coal retain their home cultures. While our model abstracts from many developments of the period that ought to contribute to cultural change, this exercise suggests that the mechanisms we do focus on—economic change driving migration and cultural sorting—plausibly account for the specific patterns of identity change we observe in our data.

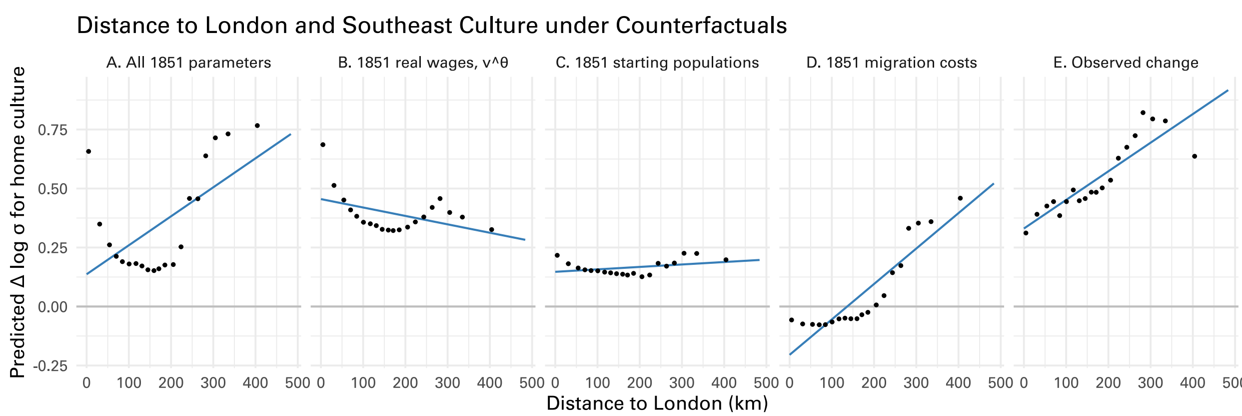

Figure 10: Distance to London, coal, and the changing popularity of the Southeast English culture over the late 19th century

This figure shows the predicted change in the log share allocated names most associated with the Southeast culture, across different counterfactual scenarios. The y axis is the observed log share minus the counterfactual log share, the x axis distance from the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B, only 1851 real wages, C, 1851 migration costs, and D, 1851 starting populations. Panel E shows the observed change going from the 1841–1860 generation to the 1860–1895 generation.

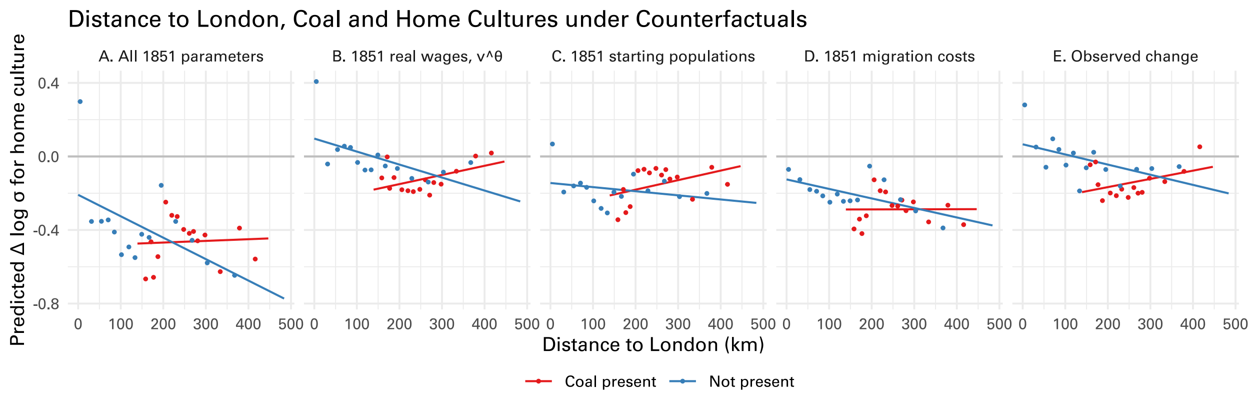

Figure 11: Distance to London, coal, and the changing popularity of home cultures over the late 19th century

This figure shows the predicted change in the log share allocated names most associated with the home culture, across different counterfactual scenarios, subset by whether the district contains a coal deposit. The y axis is the observed log share minus the counterfactual log share, the x axis distance average distance to other districts, weighted by 1851 population. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B, only 1851 real wages, C, 1851 migration costs, and D, 1851 starting populations. Panel E shows the observed change going from the 1841–1860 generation to the 1860–1895 generation.

7.2 *Understanding the Contribution of Different Economic Mechanisms to Identity Change*

7.2.1 The Role of Industrial Change

Having established that changes to the model's economic fundamentals contributed to observed patterns of cultural change, we next separate out the contributions of different components of economic change. In the first of these exercises, we fix cultural parameters, population, and migration costs at 1911 levels, and solve for the counterfactual equilibrium under 1851 real wages. Effectively, we compare England and Wales in 1911 to a hypothetical country where everything is the same, but real wages remained at their 1851 levels. This simulation examines the cultural effects of economic incentives to migrate to different districts provided by growth in industrial activity.

The results indicate that the spatial distribution of industrial change was the main contributor to the first stylized fact we documented, the spread of the culture of the Southeast. The second row of Table 6 indicates that going from 1911 to 1851 real wages generates a 35% decrease in the popularity of the Southeast culture, larger than the entire decrease predicted by the counterfactual using all three sets of fundamentals. The predicted change in cultural shares from this real wage counterfactual explains around a quarter of the raw variation in observed changes, much of which is attributable to variation in which cultures grew and declined (Table B.8).

Table 6: Counterfactual estimates

	Home culture pop.		S E Culture pop.		Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	fixing pop.
All 1851 parameters	3.193	-0.611	-31.448	-28.323	-35.670	-30.983
1851 real wages, v_d^θ	-16.692	-16.692	-35.070	-35.070	-1.562	-1.562
1851 starting populations	-9.272	-7.048	-22.382	-14.625	-0.782	0.561
1851 migration costs	16.117	16.117	-1.058	-1.058	-31.286	-31.286

This table shows percentage changes in the population choosing each district's home culture, the population choosing the Southeast English culture, and the share migrating from their district of birth. In each group, the first column gives the total percentage change in the quantity of interest under the counterfactual relative to the observed outcome, the second gives the percentage change implied by the share in each location choosing it, fixing origin populations at the observed level. To fix populations, we calculate the weighted average of the quantity of interest in the counterfactual over locations weighted by realized starting population, and dividing that by the weighted average of the realized quantity of interest, weighted by realized starting population. The first row replaces destination real wages, migration costs and starting populations with their estimated 1851 values, the second replaces only destination real wages, the third replaces starting populations, and the fourth migration costs with those estimated from 1851 data.

Why did the culture of the Southeast win out in particular? In our model, economic growth increases the popularity of a culture if it draws migrants to regions where that culture is especially popular. The left panel of Figure A.6 plots the counterfactual change in cultural populations against the change in real wage, weighted by cultural amenities. Wage growth between 1851 and 1911 was highest in regions with high cultural amenities for the Southeast, as well as for other cultures that grew such as the Northeast of England.

Industrial change does not explain the larger increase in the prevalence of the Southeast culture in the peripheries (Fact 2). Figure 10B shows the relationship between distance from London and the change in the popularity of the Southeast culture under counterfactual changes to the real wage (note that the figure examines the observed data minus the counterfactual, while Table 6 examines the opposite). The change in the popularity of the Southeast culture does not closely match any clear center-periphery pattern. Instead of the observed larger increases further from London, this simulation predicts the largest increases in the Southeast culture in places very close to London. This predicted pattern reflects relative economic growth near London that decreases out-migration from the Southeast—bolstering the Southeast culture there, and in coal-producing districts far from London, which decreases migration to the Southeast and pushes against assimilation into the Southeast. These mechanisms can be seen in Figure A.7 which plots the change in real wages against distance from London, and Figures A.8B which plot counterfactual changes in migration to the Southeast under 1851 parameters.

Changing industrial activity does instead contribute to Fact 3, the heterogeneous decline of local cultures. Figure 11B shows that in the real wage counterfactual, non-coal producing locations further from the center experience larger declines in the popularity of the home culture, while the opposite pattern is true for coal-producing locations. Indeed, it is this real wage counterfactual that most closely matches the observed data (reproduced in Figure 11E for ease of comparison). Further from the center, industrialization deters out-migration, preserving local culture: coal-producing locations experience lower rates of migration to other clusters (Figure A.10B).

Local industrialization—in both our model and data—served to preserve local cultures, but only on the periphery. Table B.9 shows that wage growth was associated with the growth of the home culture, but only in districts sufficiently far from London. This spatial pattern is reproduced in model counterfactuals that change wages. Local wage growth reduces the share migrating to other cultural regions, and has larger effects on out-migration in districts further from London (Table B.10).

7.2.2 The Role of Origin Population Changes

Next, we examine how the changing birthplaces of migrants influence outcomes, holding fixed the other economic fundamentals. Changing origin populations complement the effects of changing real wages on cultural choices. The third row of Table 6 indicates that going from 1911 to 1851 origin populations decreases the share choosing the Southeast culture by 15%, after netting out the mechanical effect on total populations from allocating population to places where people are more or less likely to choose a given culture. The logic for this effect is that population growth between 1851 and 1911 was larger in areas with higher cultural transmission tastes for the Southeast. More parents choosing to assign their children the Southeast culture in turn creates more migrants belonging to the Southeast culture and larger destination populations for the Southeast culture, which feed into the cultural choices of other parents. The right panel of Figure A.6 shows that for the Southeast and other cultures that grew over this period, population growth tended to occur in places with higher cultural transmission tastes.

As with wage changes, changing origin populations increased the prevalence of the Southeast culture everywhere (not specifically on the peripheries, see Figure 10C), and helped preserve coal-producing home cultures (Figure 11C). Population growth was highest in coal-producing districts and those close to London (Figure A.7, right panel). By influencing destination populations and thus others' migration decisions, population growth served to decrease out-migration from these kinds of areas (Figure A.10C).

7.2.3 The Role of Changing Migration Costs

Finally, we fix other economic fundamentals and solve for the counterfactual equilibrium under 1851 migration costs. Changes in migration costs are the main driver of the observed core-periphery differences in cultural change across districts (Fact 2). Falling migration costs during the period drove large increases in the overall level of migration. Table 6 indicates that going from 1911 to 1851 migration costs, holding all else fixed, results in a 31% decrease in the share of the population migrating outside the district of birth. This in turn has implications for spatial patterns of identity change. Figure 10D shows that falling migration costs increase the popularity of the Southeast culture in districts further from London. Lower migration costs imply more migration over longer distances: migration to and from the Southeast increases most in places further from the Southeast (Figures A.8D and A.9D). The same mechanism accounts for the larger observed declines in local cultures in districts further from the country's center (Figure 11D).

Lower migration costs contribute to the overall loss of home culture in districts with access to coal. Coal-producing districts experience higher rates of out-migration in this counterfactual (Figure A.10D), an effect mainly driven by their proximity to other high-real wage locations.[14]

Changing migration costs also contribute little to the overall change in popularity of the Southeast (Fact 1). Lower migration costs increase the extent of migration, but have less of an effect than changing real wages on the choice of migration destinations. In addition, falling migration costs have offsetting effects on cultural choices: while locations far from London are predicted increases in the popularity of the Southeast culture due to increased migration to and from the Southeast, locations near London are predicted small decreases in the popularity of the Southeast culture due to increased migration to and from other clusters.

8 USING THE MODEL TO UNDERSTAND HOW CULTURAL CHOICES RESPOND TO LOCAL ECONOMIC CHANGE

Does industrialization support or undermine local cultures? Our analysis thus far suggests that economic growth has offsetting effects on culture: discouraging out-migration but encouraging in-migration. In this section we use the model to examine how shocks to different locations lead to changes in the popularity of different cultures. This approach is distinct from that of the previous section, which examined how realized economic changes between 1851 and 1911 affected cultural choices. Studying the observed change only lets us learn how one particular bundle of economic changes affected culture in all locations.

[14] Rates of in-migration from other clusters are similar, as shown in Figure A.11.

For each district, we run a separate simulation in which we increase the real wage v_d^θ by 1% in that district, and record the change in cultural choices in that district, which we convert into an elasticity. This model-implied elasticity should indicate which cultures grow or decline in a district in response to economic growth in that district. We validate that claim in Appendix Section G, Table G.1, where we regress changes in cultural choices at the district-culture level against district-level economic growth. The coefficient on the interaction between growth and the model-implied elasticity is positive, indicating that cultures grew in response to economic growth where the elasticity would predict such growth.

How cultural choices respond to economic change depends upon geography. As shown in Figure 12, in locations on the periphery, an increase in v_d^θ bolsters the local culture and reduces the appeal of the Southeast culture. In the center, local industrialization shores up the culture of the Southeast, in some cases at the expense of the home culture. Figure 13 plots these elasticities against average distance to other districts and provides more direct evidence of this core-periphery pattern. As one goes further from the center, shocks to real wages increase the home culture and decrease the Southeast culture. In Table G.1, we find that the tendency of wage growth in districts further from London to preserve the home culture is attributable to the greater model-implied elasticity for the home culture further from London.

What explains the variation in how cultural choices respond to local economic change? While we cannot derive closed-form expressions for the general equilibrium elasticities, we can analyze the elasticities predicted by the model holding some of the general equilibrium components fixed. Log-differentiating equation (3) gives the following expression for the elasticity of σ_o^k to v_o^θ:

$$\frac{\partial \ln \sigma_o^k}{\partial \ln v_o^\theta} = \varphi \frac{\partial \ln \Omega_o^k}{\partial \ln v_o^\theta} - \varphi \sum_{l=1}^{K} \sigma_o^l \frac{\partial \ln \Omega_o^l}{\partial \ln v_o^\theta}.$$

Note that the term on the right is a summation over all different cultures, and so will feature in elasticities of cultural choice to economic change regardless of the culture in question. Within-location variation in how individuals embrace different cultures in response to economic change will depend only on the left term, the elasticity of Ω_o^k to v_o^θ. The real wage v_o^θ enters Ω_o^k both directly, and by affecting the number of members of culture k in o through migration. Taking the log partial derivative of Ω_o^k with respect to v_o^θ, holding m_o^k constant, gives the following expression:

$$\frac{\partial \ln \Omega_o^k}{\partial \ln v_o^\theta} = \frac{\left(v_o \delta_{oo} \xi_o^k\right)^\theta \left(m_o^k\right)^\alpha}{\Omega_o^k}. \tag{6}$$

This partial equilibrium elasticity is equivalent to the share of people of culture k in location

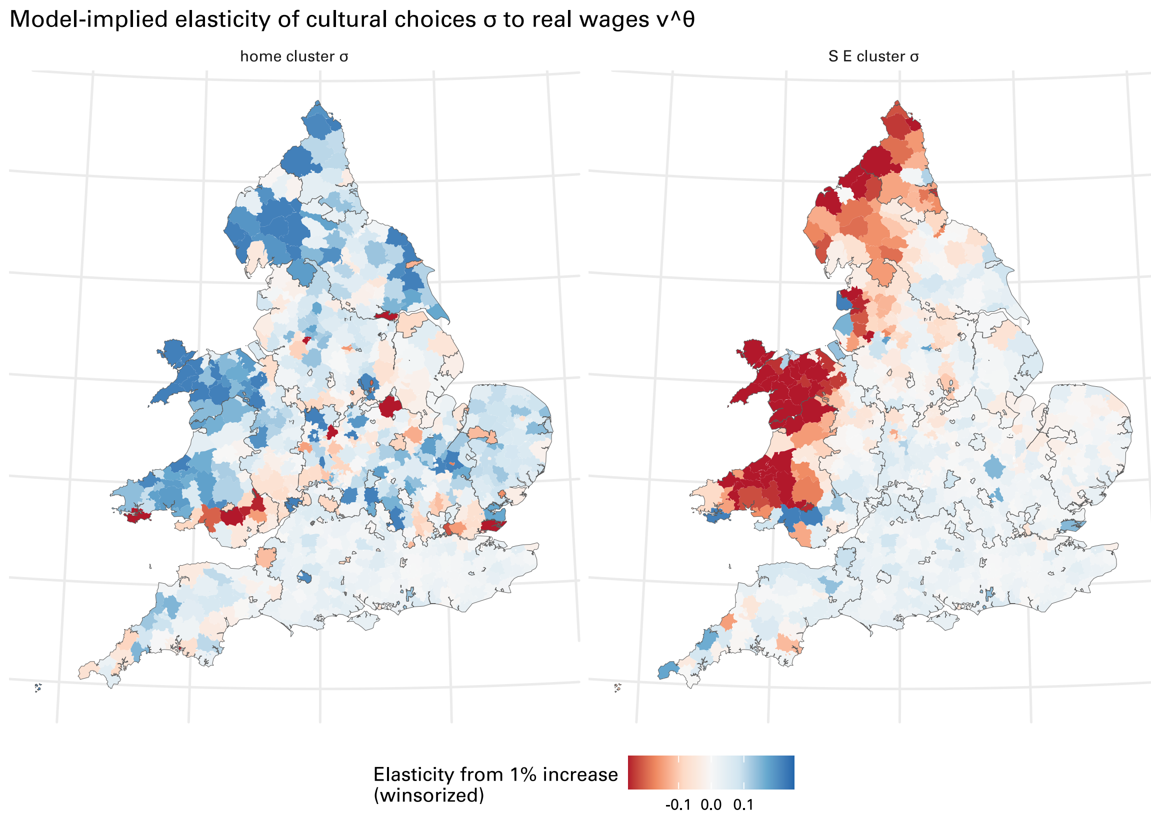

Figure 12: Elasticities of cultural choices to real wages

We run one counterfactual simulation for each district in which we increase the real wage v_d^θ by 1%, and solve for the new equilibrium cultural choices σ in each location. The map on the left shows the elasticity implied by this 1% shock of σ for each location's home culture; the map on the right shows the elasticity of σ for the Southeast culture.

o who choose to remain in location o. We calculate this quantity using the gravity fixed effects (which provide $\left(v_o\xi_o^k\right)^\theta\left(m_o^k\right)^\alpha$), estimated migration costs, and calculated values of Ω_o^k from Section 5. Conditional on district fixed effects, this measure is very strongly correlated with the general equilibrium elasticities (Table G.2), with an R^2 of 0.92.

Equation (6) helps us say more about which cultures in which locations will grow in response to local economic growth, and why. The culture-specific components of the numerator capture exogenous (ξ_o^k) and endogenous (m_o^k) preferences of people of culture k for that location. The denominator (Ω_o^k) captures these culture-specific preferences for other locations, weighted by real wages and migration costs. A positive shock to the real wage in location o influences people's choices for culture k more than other cultures if there are stronger non-economic benefits to being of culture k if one lives in o than in other locations, relative to that locational tradeoff for other cultures.

The lower rates of migration to and from peripheral regions account for the core-periphery

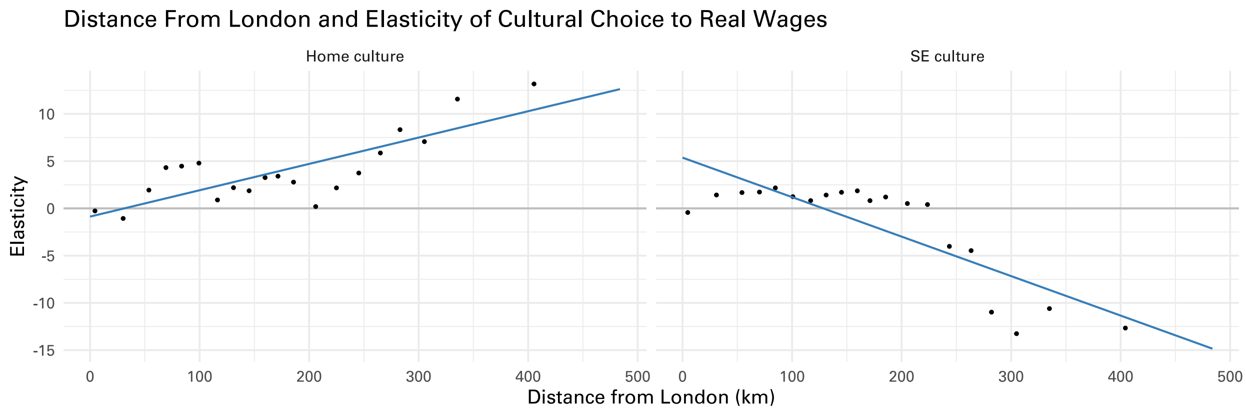

Figure 13: Elasticity of cultural choices to real wages, relative to distance to London

This figure plots the general equilibrium elasticity of the home culture (left) and Southeast English culture (right) to the district real wage, against distance to London.

pattern in home culture elasticities. Figure 14A plots the partial equilibrium elasticity for the home and other cultures, against distance from London. Going further from London, the home culture elasticity increases more, indicating that a positive shock to wages increases the proportion choosing the home culture. Panels B and C decompose the contributions of culture-specific populations (m_d^k) and cultural amenities (ξ_d^k), by calculating the equivalent of the partial equilibrium elasticity excluding cultural amenities (in B) or population (in C). These figures show that culture-specific patterns of migration, not differences in cultural amenities, account for the core-periphery pattern in home culture elasticities. In peripheral regions, distance-related migration costs lead to less out-migration by members of the home culture and less in-migration by other cultures which serves to reinforce the home culture. Examining the partial equilibrium elasticity for the Southeast culture relative to other cultures in Figure 15 reveals an essentially flipped pattern: locations further from London attract fewer members of the Southeast culture and so have weaker elasticities for that culture relative to others.

9 Summary of Additional Results and Robustness Checks

This section reviews a number of additional analyses that extend or reinforce our findings. First, in Appendix H.1 we examine how cultural homogenization influences economic output. Total homogenization into the culture of the Southeast, calculated by setting the cultural transmission tastes of all other cultures to zero, increases the average v_d^θ by 0.8%. The logic for this result is that cultural amenities (ξ_d^k) for the Southeast culture are correlated with higher real wages; assimilation into the Southeast culture moves workers out of low-

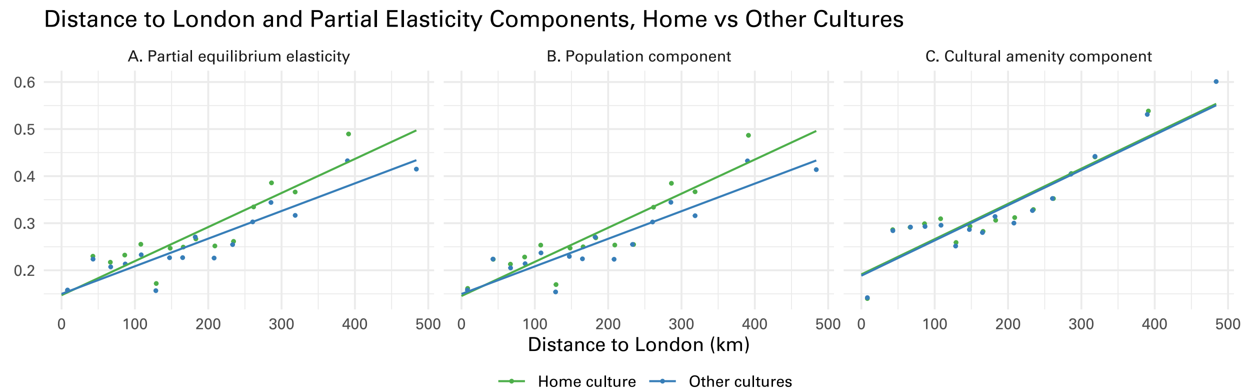

Figure 14: Decomposing the relationship between distance to London and partial equilibrium elasticities for different cultures

Panel A plots the partial equilibrium elasticity, defined in Equation (6) for the home culture and other cultures, against distance to London. The figure is a binned scatterplot with OLS overlay. The larger gap between the home and other culture values further from London indicates a stronger predicted effect of real wages on choosing the home culture further from London. Panel B plots the *population component*$_o^k := \frac{(v_o\delta_{oo})^\theta \left(m_o^k\right)^\alpha}{\sum_{d=1}^N (v_d\delta_{od})^\theta \left(m_d^k\right)^\alpha}$ against distance to London. This variable corresponds to the partial equilibrium elasticity if cultural amenities were the same everywhere and measures the extent to which the partial equilibrium elasticity is influenced by the culture's population being concentrated in the location in question. Panel C plots the *cultural amenity component*$_o^k := \frac{\left(v_o\delta_{oo}\xi_o^k\right)^\theta}{\sum_{d=1}^N \left(v_d\delta_{od}\xi_d^k\right)^\theta}$ against distance to London. This variable is the partial elasticity if destination populations did not influence migration and cultural choices and measures the extent to which the partial equilibrium elasticity is influenced by amenities for the culture being concentrated in the location in question.

wage regions with high cultural amenities for other cultures. The small magnitude of this effect makes sense given the paper's other findings. Distance-related migration costs that help preserve peripheral cultures by diminishing out-migration also keep workers in low-productivity peripheral regions, even in the absence of cultural sorting and preferences. The model already assumes that individuals migrate towards higher-wage locations, which limits the extent to which eliminating other factors can shift migration towards such locations.

Second, we address the concern that our measure of migration costs which relies on geographic distance does not account for additional factors affecting the ease of moving across locations, such as transport links. In Appendix H.2 we develop measures of transportation costs between locations following Donaldson and Hornbeck (2016). We create a network database using shapefiles for railways and shipping routes and use it to calculate least cost paths between all districts. While we do find that the implied transportation costs negatively correlate with migration flows, this association is entirely captured by geographic distance. Measures of migration costs incorporating transportation infrastructure would not improve on the distance-related measures we use in our main analyses.

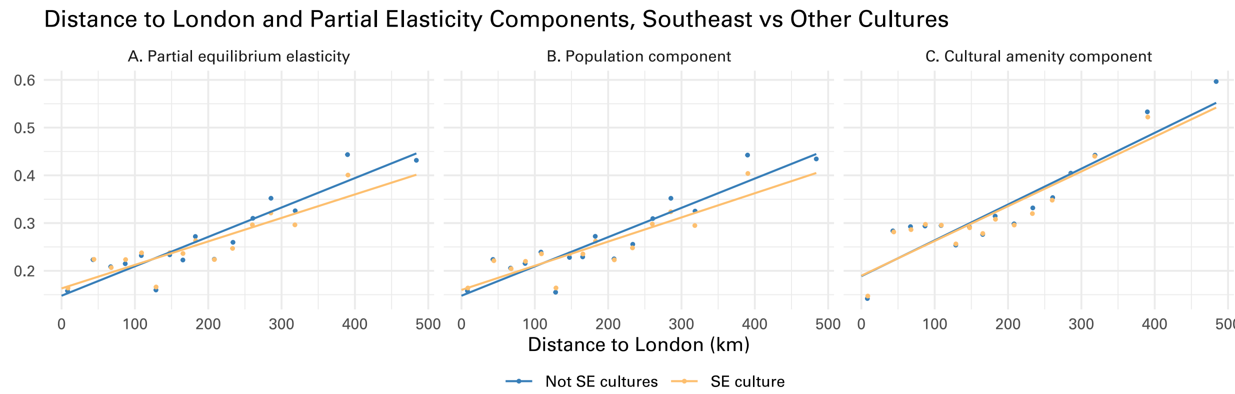

Figure 15: Decomposing the relationship between distance to London and partial equilibrium elasticities for the Southeast and other cultures

Panel A plots the partial equilibrium elasticity, defined in Equation (6) for the Southeast English culture and other cultures, against distance to London. The figure is a binned scatterplot with OLS overlay. Panel B plots the contribution of population to this pattern, as defined in Figure 14, C plots the contribution of cultural amenities, also as defined in Figure 14.

Third, in Appendix H.3 we examine whether our measures of cultural amenities actually capture local cultural wage premia. This could be the case if migrants from specific cultures have skills that are better matched to the economy of the destination, or if there is labor market discrimination against certain cultural origins. Using data on the economic status of jobs held by members of different cultures in different locations, we find a null or negative relationship between cultural amenities and status. A negative relationship would be consistent with a negative compensating differential for amenities as in canonical economic geography models (Roback, 1982; Bryan and Morten, 2019), which increases our confidence in our interpretation of ξ_d^k as a cultural amenity.

Fourth, in Appendix H.4, we discuss how accounting for emigration affects our estimation and conclusions. Over the course of the 19th century vast numbers of English and Welsh people emigrated, especially to the US, Canada, Australia, and New Zealand. We show that our theory and evidence related to internal migration is unaffected by the existence of emigration, because our data conditions on not emigrating. The potential for emigration does affect cultural choices, but in our framework it is captured by the location-by-culture cultural transmission tastes. Our counterfactual exercises must be interpreted as holding the ratio of domestic to international migration opportunities fixed.

Fifth, in Appendix H.5 we examine whether changing the homophily (α) and culture (φ) elasticities alters our results. Figures H.3 and H.4 show that the broad stylized facts our model reproduces hold at a wide range of feasible elasticities, although the magnitude of the relationship, for instance, between distance to London and the growth of the Southeast, varies considerably.

Sixth, in Appendix H.6 we explore the robustness of our results to alternative cultural clusters. We show that both the growth of the Southeast and the ability of our model to rationalize its rise across geography hold using either regions used to organize civil defense, or clusters estimated using medieval linguistic data.

10 CONCLUSION

We have examined the role of migration in changing the cultural map of 19th century industrializing England and Wales. Using rich census microdata on individuals' names and migration choices, we find that industrialization during the Second Industrial Revolution led to the decline of local cultures and a shift towards the culture of London, but also heterogeneity in the degree of cultural change. Assimilation into Southeast culture and the loss of local identities was more pronounced in the periphery. Yet no such gradient existed among districts with access to coal; coal-producing peripheries resisted the overall tendency of peripheries to be drawn closer to London, remaining relative cultural holdouts.

We develop and estimate a quantitative spatial model that allows us to understand the role that different economic components of modernization played in driving this heterogeneity. The model relies on the assumption – based on empirical observations in the data – that cultural choices and migration decisions influence one another. Parents make cultural investments in their children in anticipation of future migration opportunities and children choose migration destinations partly based on the desire to be with others who share their identity. Though these mechanisms only account for part of the many processes that drove identity change during the period we study, the model matches closely the patterns we observe in the data, pointing to the relevance of the channels we study. Counterfactual exercises help distinguish the contribution of two major economic changes occurring in Britain at the time: changing patterns of industrial activity and falling costs of migration. The results indicate that industrial change drove the spread of the Southeast culture, while falling migration costs were the main driver of the observed core-periphery gradient in cultural homogenization. By increasing migration flows over larger distances, lower migration costs brought peripheral locations closer to London. Yet industrialization in coal-producing peripheries was able to counteract this pattern, helping preserve local identities.

One central contribution of our model is to identify a crucial mechanism behind the heterogeneous effects of industrial development on identity change: industrialization leads to both higher in-migration from diverse cultures and lower out-migration of locals. Which of the two offsetting forces dominates depends on an industrial center's geographic location relative to others. While the in-migration channel may dominate in central locations, peripheral locations that industrialize experience less out- relative to in-migration, and are more likely to retain their local identity.

Insights from our study elucidate the role of labor migration for cultural change in 19th century Britain and complement existing explanations for the survival of regional differences. Most prominently, Hechter (1977) proposed a theory of economic and sociopolitical dominance of the core on the peripheries to explain the strength of local identity in Wales, Scotland and Ireland. Our model supports some of the patterns and mechanisms identified by Hechter, such as that industrialization drives overall homogenization and that decreased out-migration from Wales in response to coal-driven industrial development promoted the retention of the Welsh language (until English language policies explicitly targeted it for eradication in the early 20th century). As in Hechter, our model also emphasizes heterogeneity within the peripheries. However, while his theory predicts highest assimilation in peripheral industrial hubs, our emphasis on the interplay between labor migration and cultural choice indicates that local industrial development may work in favor, rather than against, the retention of local culture, counteracting other forces of modernization that pulled the peripheries toward the core.

How do the insights from 19th century Britain travel to other country contexts? Our framework is general enough to apply to any situation in which economic activity changes across space and people make migration decisions driven by both economic and cultural considerations. Empirically, the result that peripheral regions may be more likely to resist assimilation into a dominant culture if they develop economically could rationalize patterns we observe in other parts of Europe during the 19th and early 20th centuries. The Basque country in Spain was able to retain its distinctive identity among other reasons because the development of ore production and steel manufacturing directed local rural migrants to head to Bilbao rather than Madrid (Green, 2022, p.52). Scotland and Catalonia in Spain were additional cases of industrializing peripheral areas with distinct identities. Discouragement of out-migration may have contributed to cultural retention in those regions, while peripherality muted the effects of in-migration from different cultures into their main industrial centers.

The migration link between industrialization and cultural change also informs a broader literature on nation-building. The formation of national identities is thought to be primarily a task of the state, which homogenizes populations through coercion (Tilly, 1975; Mylonas, 2013), conscription (Weber, 1976) or education (Hobsbawm, 1992; Cinnirella and Schueler, 2018; Alesina, Giuliano and Reich, 2021). Our findings show that beyond top-down efforts to build national identities, bottom-up processes such as spatial changes in the pattern of economic activity can also play a crucial role in cultural unification. This in turn has implications for the role of the state, suggesting that less conventional tools such as industrial policy, besides their more well-documented role in restructuring economic activity (Juhász, Lane and Rodrik, Forthcoming), can be used for identity engineering and nation-building.

REFERENCES

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm and Nikolaus Wolf. 2015. "The Economics of Density: Evidence From the Berlin Wall." *Econometrica* 83(6):2127–2189.

Alesina, Alberto, Paola Giuliano and Bryony Reich. 2021. "Nation-building and education." *The Economic Journal* 131(638):2273–2303.

Allen, Treb and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *The Quarterly Journal of Economics* 129(3):1085–1140.

Allen, Treb, Costas Arkolakis and Xiangliang Li. 2020. On the Equilibrium Properties of Network Models with Heterogeneous Agents. Working Paper 27837 National Bureau of Economic Research.

Allen, Treb and Dave Donaldson. 2020. Persistence and Path Dependence in the Spatial Economy. Working Paper 28059 National Bureau of Economic Research.

Alvarez-Palau, Eduard J., Oliver Dunn, Dan Bogart, Max Satchell and Leigh Shaw-Taylor. 2019. "Historic ports and coastal sailing routes in England and Wales, 1540-1914.". **URL:** *https://reshare.ukdataservice.ac.uk/853711/*

Anderson, Benedict. 1983. *Imagined Communities: Reflections on the Origins and Spread of Nationalism*. New York and London: Verso.

Antràs, Pol and Hans-Joachim Voth. 2003. "Factor prices and productivity growth during the British industrial revolution." *Explorations in Economic History* 40(1):52–77.

Baker, Stuart G. 1994. "The Multinomial-Poisson Transformation." *Journal of the Royal Statistical Society: Series D (The Statistician)* 43(4):495–504.

Bates, Robert H. 1974. "Ethnic competition and modernization in contemporary Africa." *Comparative political studies* 6(4):457–484.

Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg and Maisy Wong. 2019. "Unity in Diversity? How Intergroup Contact Can Foster Nation Building." *American Economic Review* 109(11):3978–4025.

Bazzi, Samuel, Martin Fiszbein and Mesay Gebresilasse. 2020. "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States." *Econometrica* 88(6):2329–2368.

Bennett, Robert J., Harry Smith, Carry van Lieshout and Gill Newton. 2017. Business sectors, occupations and aggregations of census data 1851-1911. Working Paper 5 University of Cambridge, Department of Geography and Cambridge Group for the History of Population and Social Structure. Issue: 5.

Benskin, M., M. Laing, V. Karaiskos and K. Williamson. 2013. "An Electronic Version of A Linguistic Atlas of Late Mediaeval English.".
URL: *http://www.lel.ed.ac.uk/ihd/elalme/elalme.html*

Bisin, Alberto and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97(2):298–319.

Bisin, Alberto and Thierry Verdier. 2014. "Trade and Cultural Diversity." *Handbook of the Economics of Art and Culture* 2:439–484.

Bogart, Dan, Xuesheng You, Eduard J Alvarez-Palau, Max Satchell and Leigh Shaw-Taylor. 2022. "Railways, divergence, and structural change in 19th century England and Wales." *Journal of Urban Economics* 128:103390.

Borusyak, Kirill and Peter Hull. 2020. Non-Random Exposure to Exogenous Shocks: Theory and Applications. Working Paper 27845 National Bureau of Economic Research.

Bryan, Gharad and Melanie Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia." *Journal of Political Economy* 127(5):2229–2268.

Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg and Pierre-Daniel Sarte. 2018. "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy." *The Review of Economic Studies* 85(4):2042–2096.

Caliendo, Lorenzo, Luca David Opromolla, Fernando Parro and Alessandro Sforza. 2021. "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement." *Journal of Political Economy* 129(12):3491–3545.

Cinnirella, Francesco and Ruth Schueler. 2018. "Nation Building: The Role of Central Spending in Education." *Explorations in Economic History* 67:18–39.

Colley, Linda. 2005. *Britons: Forging the Nation, 1707-1837*. New Haven and London: Yale University Press.

Crouzet, François. 1982. *The Victorian Economy*. London: Methuen and Co.

Day, Joseph. 2016. "Consistent parish cross-walk files.". Work conducted as part of the ESRC-funded project 'Migration, Urbanisation and Socio-Economic Change, England and Wales 1851-1911'.

Day, Joseph. 2023. "Mapping the cultural divides of England and Wales: Did the geographies of 'Belonging'act as a brake on British Urbanisation, 1851–1911?" *Plos one* 18(5):e0286244.

Deutsch, Karl Wolfgang. 1966. *Nationalism and Social Communication: An Inquiry into the Foundations of Nationality.* Cambridge, MA: MIT Press.

Deutsch, Karl Wolfgang. 1969. *Nationalism and its Alternatives.* New York: Alfred A. Knopf.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* 108(4-5):899–934.

Donaldson, Dave and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *The Quarterly Journal of Economics* 131(2):799–858.

Fajgelbaum, Pablo and Stephen J. Redding. 2022. "Trade, Structural Transformation, and Development: Evidence from Argentina 1869–1914." *Journal of Political Economy* 130(5):1249–1318. Number: 5.

Fally, Thibault. 2015. "Structural gravity and fixed effects." *Journal of International Economics* 97(1):76–85.

Fava-Verde, Jean-François. 2018. "Victorian telegrams: the early development of the telegraphic despatch and its interplay with the letter post." *Notes and Records: the Royal Society Journal of the History of Science* 72(3):275–292.

Fouka, Vasiliki. 2019. "How do Immigrants Respond to Discrimination? The Case of Germans in the US during World War I." *American Political Science Review* 113(2):405–422.

Fouka, Vasiliki. 2020. "Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I." *The Review of Economic Studies* 87(1):204–239.

Fryer Jr, Roland G and Steven D Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *The Quarterly Journal of Economics* 119(3):767–805.

Geary, Frank and Tom Stark. 2015. "Regional GDP in the UK, 1861–1911: new estimates." *The Economic History Review* 68(1):123–144.

Gellner, Ernest. 1964. *Thought and Change.* Chicago: The University of Chicago Press.

Gellner, Ernest. 2006 [1983]. *Nations and Nationalism.* Oxford: Blackwell Publishing.

Green, Elliott. 2019. "Industrialization and ethnic change in the modern world." *Ethnic and Racial Studies* 42(2):178–197.

Green, Elliott. 2022. *Industrialization and Assimilation: Understanding Ethnic Change in the Modern World.* Cambridge University Press.

Görlach, Manfred. 1999. *English in Nineteenth-Century England: An Introduction.* Cambridge: Cambridge University Press.

Hechter, Michael. 1977. *Internal colonialism: The Celtic fringe in British national development, 1536-1966.* University of California Press.

Henneberg, J., M. Satchell, X. You, L. M. W. Shaw-Taylor, E. A. Wrigley and Michael Cobb. 2018. "1881 England, Wales and Scotland Railway Stations.".
URL: *https://reshare.ukdataservice.ac.uk/852996/*

Hobsbawm, Eric J. 1992. *Nations and Nationalism since 1780: Programme, Myth, Reality.* Cambridge University Press.

Homans, George C. 1969. "The Explanation of English Regional Differences." *Past & Present* (42):18–34.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica* 87(5):1439–1474.

John, Christopher R., David Watson, Michael R. Barnes, Costantino Pitzalis and Myles J. Lewis. 2020. "Spectrum: fast density-aware spectral clustering for single and multi-omic data." *Bioinformatics (Oxford, England)* 36(4):1159–1166.

Juhász, Réka, Nathan Lane and Dani Rodrik. Forthcoming. "The New Economics of Industrial Policy." *Annual Review Economics* .

Kandt, Jens, James A Cheshire and Paul A Longley. 2016. "Regional surnames and genetic structure in Great Britain." *Transactions of the Institute of British Geographers* 41(4):554–569.

Laitin, David D. 1994. "The Tower of Babel as a Coordination Game: Political Linguistics in Ghana." *The American Political Science Review* 88(3):622–634.

Laitin, David D. 1995. "Marginality: A Microperspective." *Rationality and Society* 7(1):31–57.

Laitin, David D. 2007. *Nations, States, and Violence*. Oxford, UK: Oxford University Press.

Lambert, Paul S., Richard L. Zijdeman, Marco H. D. Van Leeuwen, Ineke Maas and Kenneth Prandy. 2013. "The Construction of HISCAM: A Stratification Scale Based on Social Interactions for Historical Comparative Research." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 46(2):77–89.

Lazear, Edward P. 1999. "Culture and Language." *Journal of political Economy* 107(S6):S95–S126.

Marti-Henneberg, J., M. Satchell, X. You, L. M. W. Shaw-Taylor and E. A. Wrigley. 2018. "1851 England, Wales and Scotland Railway Stations.".
URL: *https://reshare.ukdataservice.ac.uk/852994/*

Marti-Henneberg, J., M. Satchell, X. You, L.M.W. Shaw-Taylor and E.A. Wrigley. 2023. "1881 England, Wales and Scotland Rail Lines.".

Marx, Karl and Friedrich Engels. 1978 [1848]. Manifesto of the Communist Party. In *The Marx-Engels Reader*, ed. Robert Tucker. New York: W. W. Norton pp. 594–618.

Maystre, Nicolas, Jacques Olivier, Mathias Thoenig and Thierry Verdier. 2014. "Product-Based Cultural Change: Is the Village Global?" *Journal of International Economics* 92(2):212–230.

McFadden, Daniel. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, ed. Paul Zarembka. New York: Academic Press pp. 105–142.

Mokyr, Joel. 2008. Accounting for the Industrial Revolution. In *The Cambridge Economic History of Modern Britain, Volume 1: Industrialisation, 1700-1860*, ed. Roderick Floud and Paul Johnson. Cambridge: Cambridge University Press pp. 1–27.

Monte, Ferdinando, Stephen J. Redding and Esteban Rossi-Hansberg. 2018. "Commuting, Migration, and Local Employment Elasticities." *American Economic Review* 108(12):3855–3890.

Morten, Melanie and Jacqueline Oliveira. 2023. The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City. Working Paper.

Mylonas, Harris. 2013. *The Politics of Nation-Building: Making Co-Nationals, Refugees, and Minorities*. Cambridge University Press.

Olivier, Jacques, Mathias Thoenig and Thierry Verdier. 2008. "Globalization and the Dynamics of Cultural Identity." *Journal of International Economics* 76(2):356–370.

Palgrave Macmillan Ltd. 2013. *International Historical Statistics*. London: Palgrave Macmillan.
URL: *https://doi.org/10.1057/978-1-137-30568-8*

Pengl, Yannick I, Philip Roessler and Valeria Rueda. 2022. "Cash crops, print technologies, and the politicization of ethnicity in Africa." *American Political Science Review* 116(1):181–199.

Porteous, J. Douglas. 1982. "Surname Geography: A Study of the Mell Family Name c. 1538-1980." *Transactions of the Institute of British Geographers* 7(4):395–418.

Rapoport, Hillel, Sulin Sardoschau and Arthur Silve. 2020. Migration and Cultural Change. Cesifo working paper no. 8547.

Redding, Stephen J. and Daniel M. Sturm. 2008. "The Costs of Remoteness: Evidence from German Division and Reunification." *American Economic Review* 98(5):1766–1797.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90(6):1257–1278. Number: 6.

Robinson, Amanda Lea. 2014. "National versus ethnic identification in Africa: Modernization, colonial legacy, and the origins of territorial nationalism." *World Politics* 66(4):709–746.

Satchell, A.E.M., P.M.K. Kitson, G.H. Newton, L. Shaw-Taylor and E.A. Wrigley. 2016. "1851 England and Wales Census parishes, townships and places.".

Satchell, M., E.A. Wrigley, L.M.W. Shaw-Taylor, X. You and J. Henneberg. 2023. "1851 England, Wales and Scotland Rail Lines.".

Schwartz, Robert M. 2023. The Internet of the Nineteenth Century: Railways and the Postal Service in France and Great Britain, 1830–1914. In *Creative Ways to apply Historical GIS : Promoting Research and Teaching about Europe*, ed. Jordi Martí-Henneberg. Cham: Springer International Publishing pp. 97–114.

Schürer, Kevin and Edward Higgs. 2014. "Integrated Census Microdata (I-CeM), 1851-1911.". SN: 7481.
URL: *http://doi.org/10.5255/UKDA-SN-7481-1*

Shaw-Taylor, Leigh, Xuesheng You, D Bogart and AEM Satchell. 2018. "The development of the railway network in Britain 1825–1911." *The Online Historical Atlas of Transport, Urbanization and Economic Development in England and Wales c. 1680–1911* .

Silva, J. M. C. Santos and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics* 88(4):641–658.

Southall, Humphrey R. and Paul Ell. 2022. "Great Britain Historical Database: Census Data: Religion Statistics, 1851 [data collection] 2nd Edition.".

Tilly, Charles. 1975. Reflections on the History of European State-Making. Princeton, NJ: Princeton University Press pp. 3–83.

Tombe, Trevor and Xiaodong Zhu. 2019. "Trade, Migration, and Productivity: A Quantitative Analysis of China." *American Economic Review* 109(5):1843–1872.

Waters, Mary C. 1990. *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.

Weber, Eugen. 1976. *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*. Stanford University Press.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second edition ed. Cambridge MA: MIT Press.

Online Appendix

Table of Contents

A additional figures

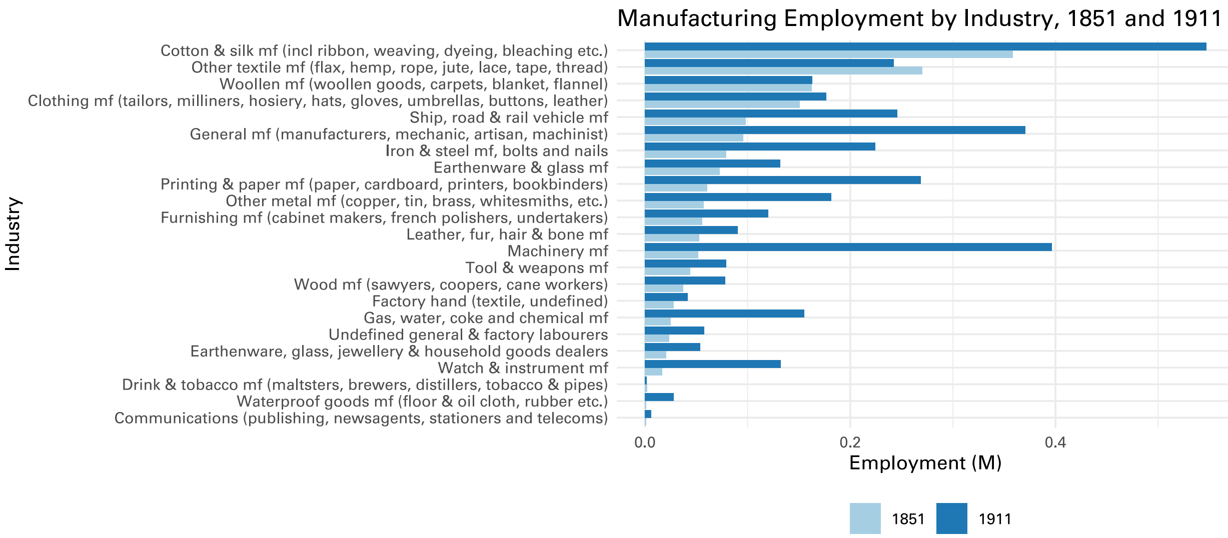

Figure A.1: Employment by manufacturing industry

This figure plots employment in different manufacturing industries in the 1851 and 1911 censuses, calculated using census microdata from Schürer and Higgs (2014) and industry categories from Bennett et al. (2017).

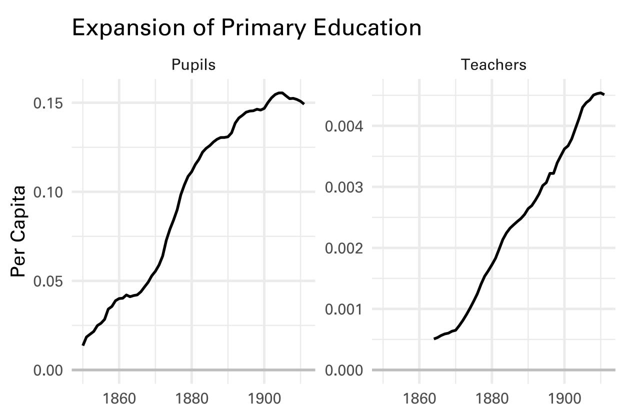

Figure A.2: Primary school teachers and pupils per capita

This figure plots the number of primary school teachers and primary school pupils, divided by total population. Source: Palgrave Macmillan Ltd (2013)

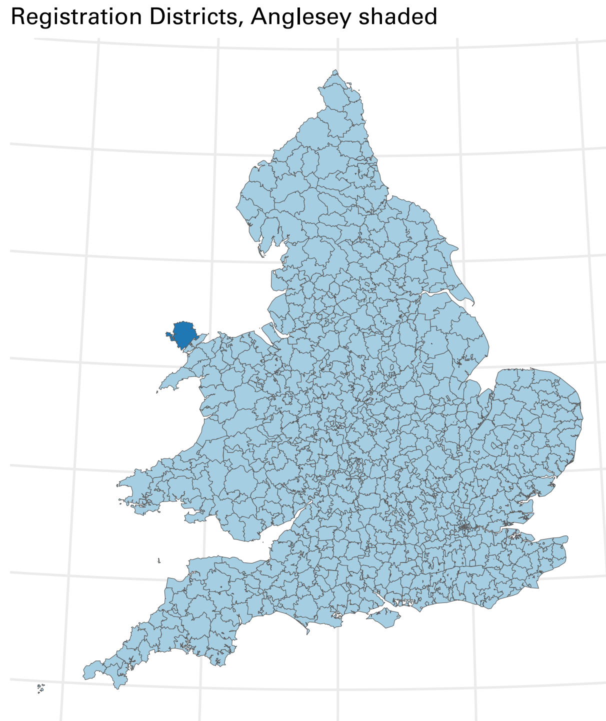

Figure A.3: Map of registration districts highlighting the location of Anglesey

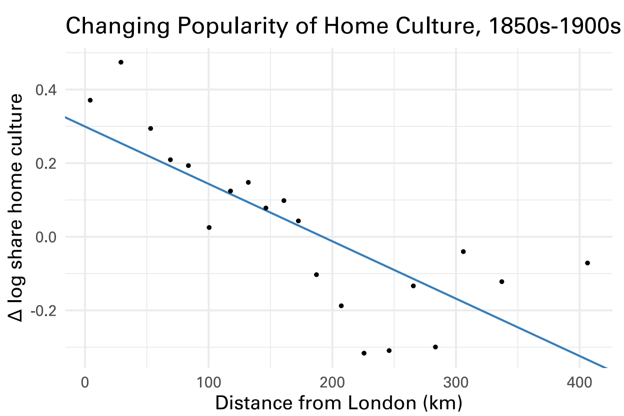

Figure A.4: Change in popularity of the home culture plotted against distance to London

This figure shows the change in the log share allocated names most associated with the home culture comparing those born between 1851–1860 and 1901–1910, plotted against the distance from the district centroid to the City of London.

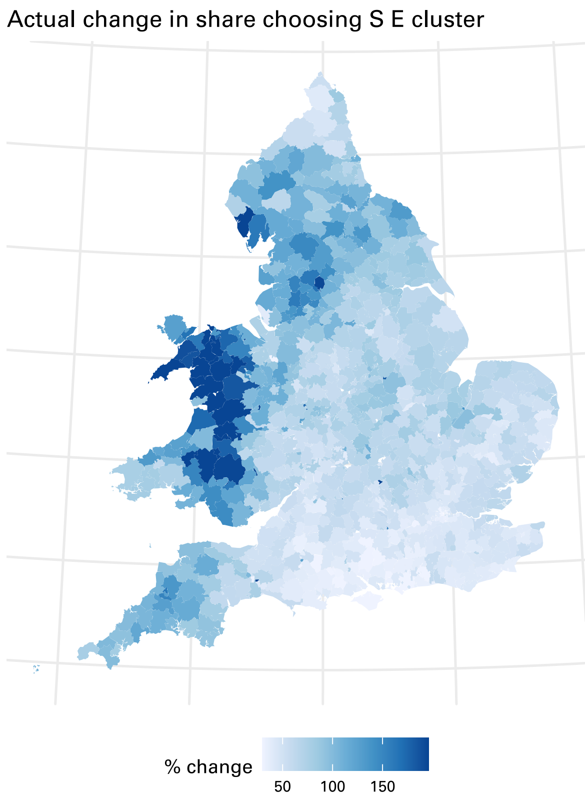

Figure A.5: Change in log share choosing the Southeast England cluster

Left panel is observed change between those born 1841–1860 and those born 1861–1895, right panel is change between counterfactual using 1851 estimates of migration costs, population, and destination utilities, and observed values for those born 1861–1895

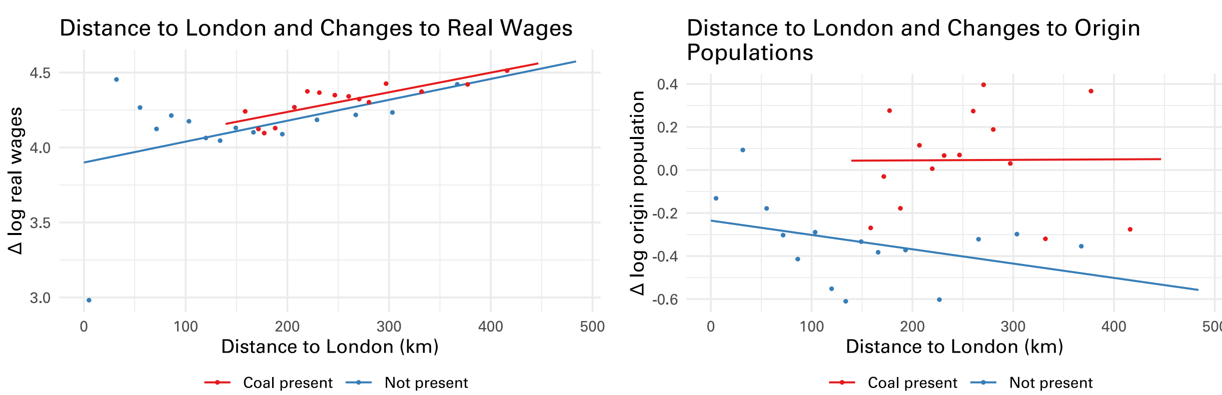

Figure A.7: Change in real wage and origin populations between 1851 and 1911, binned averages plotted against distance from London, subset by whether the district contains coal

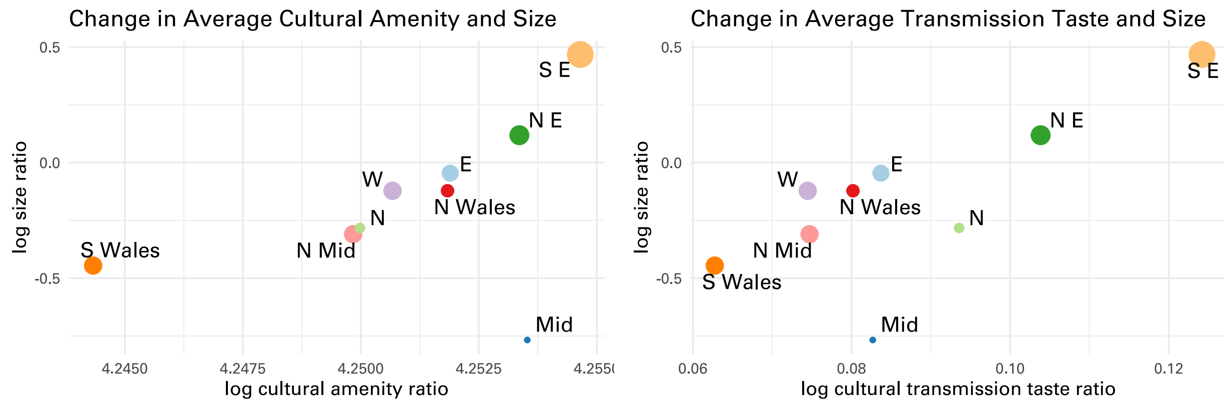

Figure A.6: Changes in average cultural amenities and transmission tastes and predicted changes in culture sizes

In both figures, the y axis is the log of the size of the cluster in the observed data, divided by the size of the cluster in a counterfactual with 1851 real wages and starting populations. To net out the mechanical effects of population growth, we use the predicted cultural choices (σ) from this counterfactual and multiply by observed starting populations. The x axis in the left panel is $\log \text{cultural amenity ratio}^k = \log\left(\frac{\sum_{d=1}^{N}\left(\xi_d^k v_{d,1911}\right)^{\theta}}{\sum_{d=1}^{N}\left(\xi_d^k v_{d,1851}\right)^{\theta}}\right)$. The x axis in the right panel is $\log \text{cultural transmission taste ratio}^k = \log\left(\frac{\sum_{o=1}^{N} n_{o,1911}\psi_o^k}{\sum_{o=1}^{N} n_{o,1851}\psi_o^k}\right)$. The size of the dots corresponds to the size of the cluster in the observed data.

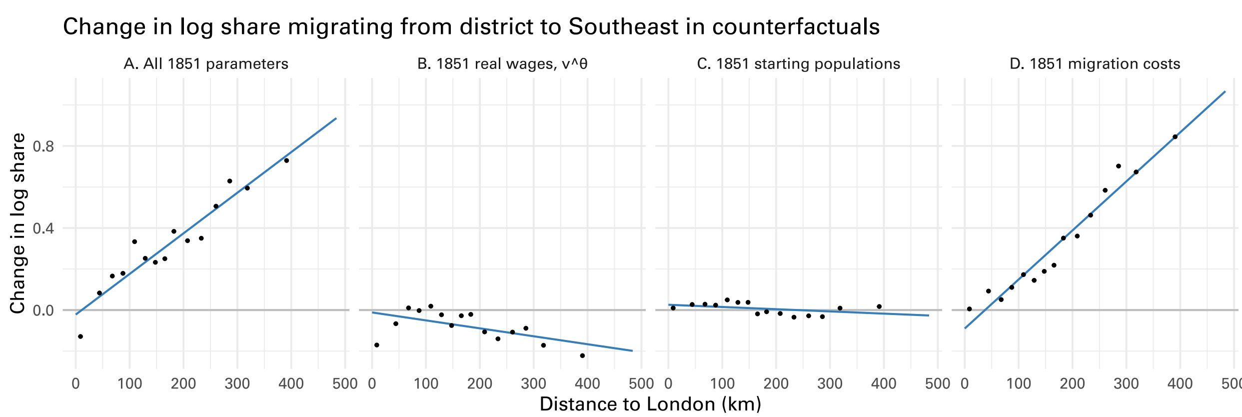

Figure A.8: Counterfactual changes in migration to districts allocated to the Southeast England cluster

The x axis is distance to London, the y axis is observed log share migrating from district in question to districts allocated to the Southeast, minus log share under counterfactual scenario.

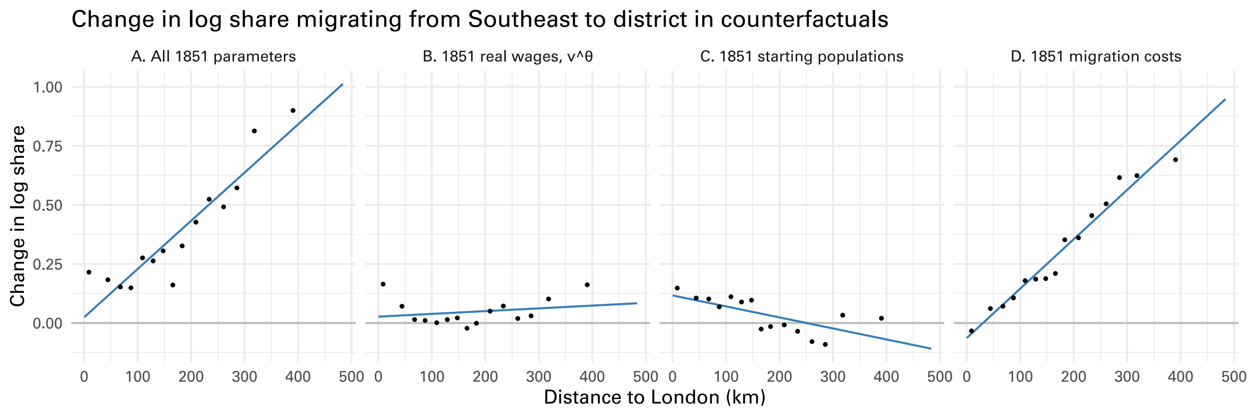

Figure A.9: Counterfactual changes in migration from the Southeast England cluster to districts

The x axis is distance to London, the y axis is observed log share migrating to district in question from districts allocated to the Southeast, minus log share under counterfactual scenario.

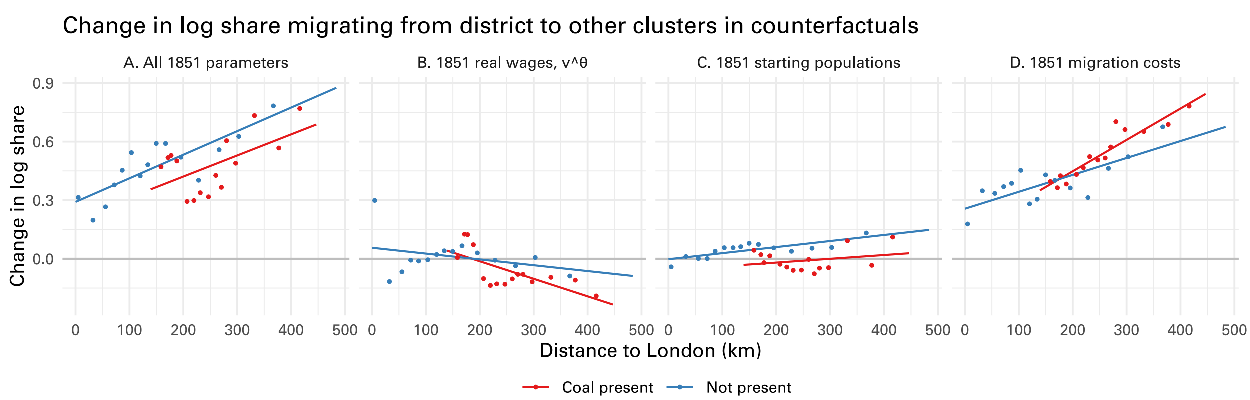

Figure A.10: Counterfactual changes in migration from the district out of the home cluster

The x axis is distance to London, the y axis is observed log share migrating from the district to districts not in the home cluster, minus log share under counterfactual scenario.

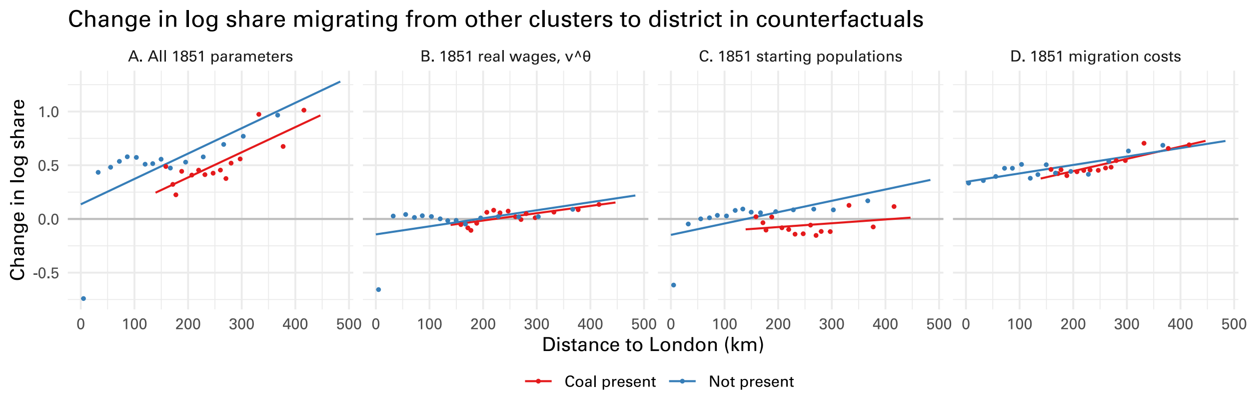

Figure A.11: Counterfactual changes in migration from non-home clusters into district

The x axis is distance to London, the y axis is observed log share migrating to the district from districts not in the home cluster, minus log share under counterfactual scenario.

B Additional tables

	ln migrants		
	(1)	(2)	(3)
different cluster	−1.861	−0.226	−0.083
	(0.074)	(0.050)	(0.033)
ln distance		−1.532	−1.475
		(0.083)	(0.075)
county-pair FE			x
N	273400	273400	273400
pseudo R^2	0.352	0.701	0.767

This table shows the results of regressions of log migration flows between registration districts in 1851 against an indicator for whether the origin and destination district are in different clusters, and log distance between the subdistrict centroids. The unit of analysis is the origin-destination district pair, restricted to pairs where the origin does not equal the destination. Models are estimated by poisson pseudo-maximum likelihood. All models have fixed effects for the origin and destination district. Model (3) adds fixed effects for each pair of origin and destination counties. Standard errors clustered by origin and destination district in parentheses.

Table B.1: People in 1851 tended to migrate within-cluster

	probability of marriage ×1000			
	(1)	(2)	(3)	(4)
different cluster	−0.376	−0.102	−1.141	−0.924
	(0.028)	(0.010)	(0.065)	(0.056)
same district		0.931		1.036
		(0.062)		(0.071)
parish FE	x	x		
cluster-by-parish FE			x	x
DV mean	1.12	1.12	1.12	1.12
N	321382	321382	321382	321382
R^2	0.531	0.539	0.535	0.545

This table shows the results of regressions of the probability of a given man-woman couple resident in the same parish being married (multiplied by 1000) against an indicator for whether both were born in the same cluster. The unit of analysis is the parish-man cluster-woman cluster-same district indicator, and observations are weighted by the number of potential couples in the unit—this is equivalent to a regression at the individual parish-man-woman level in which the dependent variable is a binary indicator of whether they are married. Models (1) and (2) have parish of residence fixed effects, (3) and (4) add fixed effects for the residence parish interacted with the man and woman's birth cluster. (2) and (4) also control for an indicator that both man and woman were born in the same registration district. Standard errors clustered by parish in parentheses.

Table B.2: People in 1851 tended to marry within-cluster

Table B.3: First names with highest name scores by cluster

			Frequency 1841–1860		
Cluster	Metaphone	Example Name	Cluster	Total	Name Score
E	ELFN	Elvina	51	142	0.918
E	HKR	Hagar	25	89	0.886
E	MHL	Mahala	108	449	0.863
E	TMRS	Damaris	17	72	0.860
E	EFRT	Everett	14	67	0.840
E	ETKR	Edgar	241	1190	0.835
E	LRN	Lorna	14	70	0.832
E	OSBRN	Osborne	12	61	0.830
E	BRNBS	Barnabas	12	62	0.827
E	BLNT	Belinda	27	143	0.822
Mid	IX	Isiah	48	86	0.842

Mid	KMFRT	Comfort	37	70	0.823
Mid	RSHN	Rosehannah	53	109	0.798
Mid	IS	Isaiah	218	499	0.767
Mid	SLP	Zilpah	35	81	0.762
Mid	XTRK	Shadrack	53	125	0.755
Mid	SRHN	Sarahann	207	530	0.731
Mid	WLFB	Willoughby	31	79	0.727
Mid	NMN	Newman	29	76	0.726
Mid	TRSL	Drusilla	98	263	0.716
N	TKSN	Dixon	28	63	0.909
N	RXRTSN	Richardson	25	58	0.907
N	JKSN	Jackson	41	107	0.887
N	0MPSN	Thompson	61	162	0.887
N	RBSN	Robson	21	58	0.880
N	WTSN	Watson	67	185	0.879
N	RBNSN	Robinson	95	267	0.877
N	MRMTK	Marmaduke	27	78	0.873
N	BRYN	Bryan	29	91	0.859
N	WLKNSN	Wilkinson	25	84	0.848
N E	K0BRT	Cuthbert	96	226	0.961
N E	RBSN	Robson	21	58	0.950
N E	TR0	Dorothy	696	2449	0.930
N E	LNSLT	Lancelot	50	180	0.928
N E	BRBR	Barbara	406	1470	0.927
N E	ANBL	Annabella	41	153	0.924
N E	RLF	Ralph	578	2548	0.907
N E	FSTR	Foster	20	91	0.904
N E	0MSN	Thomasina	66	315	0.899
N E	ISBL	Isabella	3190	15396	0.897
N Mid	STKLF	Sutcliffe	67	70	0.986
N Mid	HWR0	Howarth	51	55	0.977
N Mid	KRNWT	Greenwood	122	132	0.973
N Mid	BT	Betty	1816	2118	0.949
N Mid	SKR	Squire	326	409	0.923

N Mid	HRTL	Hartley	160	201	0.923
N Mid	RFT	Wright	263	333	0.921
N Mid	NNS	Nancy	2043	2948	0.874
N Mid	HMLT	Hamlet	57	84	0.869
N Mid	LSTR	Lester	80	119	0.861
N Wales	KRF0	Griffith	480	803	0.973
N Wales	OWN	Owen	889	1869	0.956
N Wales	KWN	Gwen	207	460	0.952
N Wales	HF	Hugh	968	2835	0.926
N Wales	IXML	Ishmael	22	72	0.914
N Wales	HMFR	Humphrey	108	376	0.907
N Wales	EFN	Evan	665	3028	0.872
N Wales	WTKN	Watkin	26	124	0.865
N Wales	LWLN	Llewellyn	63	312	0.860
N Wales	KMR	Gomer	11	59	0.847
S E	TRTN	Trayton	56	69	0.919
S E	FLTLF	Philadelphia	93	126	0.882
S E	LPLT	Leopold	38	56	0.848
S E	WR0	Worthy	39	58	0.844
S E	ERL	Earl	30	51	0.791
S E	AKSTS	Augustus	445	776	0.780
S E	HSTR	Hester	581	1023	0.776
S E	ATLFS	Adolphus	121	214	0.774
S E	BRTRM	Bertram	33	59	0.770
S E	TKLS	Douglas	134	242	0.767
S Wales	KWNLN	Gwenllian	234	247	0.998
S Wales	JNKN	Jenkin	201	216	0.997
S Wales	HWL	Howell	140	170	0.991
S Wales	MRKN	Morgan	529	649	0.990
S Wales	EFN	Evan	2017	3028	0.979
S Wales	KMR	Gomer	38	59	0.977
S Wales	WTKN	Watkin	71	124	0.969
S Wales	ELFR	Elvira	50	97	0.960
S Wales	LWLN	Llewellyn	150	312	0.955

S Wales	KWN	Gwen	214	460	0.953
W	LFT	Loveday	62	96	0.966
W	WLMT	Wilmot	42	101	0.921
W	BS	Bessie	957	2706	0.898
W	NXLS	Nicholas	342	1073	0.883
W	0MSN	Thomasina	100	315	0.883
W	TRFN	Tryphena	68	246	0.861
W	LSNT	Lucinda	28	105	0.856
W	XRT	Charity	91	349	0.850
W	NN	Nina	120	519	0.829
W	KRS	Grace	1077	4728	0.826

	Speaks Welsh		Speaks English	
	(1)	(2)	(3)	(4)
North Wales score	0.355	0.060	−0.178	−0.031
	(0.087)	(0.017)	(0.057)	(0.010)
South Wales score	0.498	0.266	−0.019	−0.020
	(0.048)	(0.035)	(0.022)	(0.005)
Southeast England score	−0.543	−0.246	0.104	0.040
	(0.021)	(0.033)	(0.015)	(0.005)
Birthplace FE		x		x
N	97932	97932	97932	97932
R^2	0.146	0.507	0.023	0.171

This table presents evidence of the relationship between region-specific name scores and speaking Welsh or English in the 1911 census. The language question was only asked in Wales and the Isle of Mann. We exclude those recorded as speaking neither language. Data is aggregated to the name-birthplace district level, observations are weighted by the number of individuals in each cell, to give the same coefficients as an individual-level regression. In (1) and (2) the dependent variable is an indicator that the individual speaks Welsh, in (3) and (4) and indicator that they speak English. The independent variables are the name scores for the individual's name and the North Wales, South Wales and Southeast England cultural regions. Models (2) and (4) add fixed effects for the district of birth, (1) and (3) include and intercept. Standard errors clustered by district of birth in parentheses.

Table B.4: Relationship between name scores and language

	Δ ln S E	Δ ln home culture		Δ ln S E	Δ ln home culture	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to London (100km)	0.186	−0.168	−0.169	0.202	−0.156	−0.157
	(0.008)	(0.016)	(0.018)	(0.009)	(0.019)	(0.027)
Contains coal			−0.607			−0.379
			(0.140)			(0.101)
Distance to London × coal			0.194			0.118
			(0.057)			(0.043)
Weighted	x	x	x			
N	827	798	798	827	798	798
R^2	0.545	0.195	0.239	0.408	0.103	0.114

This table presents evidence of the relationship between distance to London, coal, and the strength of the Southeast English and home cultures, using data on naming for the 1851–1860 and 1901–1910 cohorts. Regressions are at the district-level. The dependent variable in (1) and (4) is the change in the log share allocated names most associated with the Southeast English cluster, between those born 1851–1860 and 1901–1910, in (2), (3), (5) and (6) it is the change in the log share allocated names most associated with the home culture. Coal presence is an indicator that the district contains coal deposits, distance to London is distance from the district centroid to the City of London, in hundreds of kilometers. Models (1)–(3) are weighted by the number allocated names in the 1851–1860 generation, (4)–(6) are unweighted. (1) and (3) correspond to the left and right hand panels of Figure 5. All models include an intercept. Robust standard errors in parentheses.

Table B.5: Relationship between coal, distance to London, and changing popularity of cultures

	Δ ln share home culture	
	(1)	(2)
Δ ln out-migrants / pop	−0.281	
	(0.061)	
Δ ln in-migrants / pop		−0.066
		(0.026)
N	783	782
R^2	0.062	0.027

This table presents the results of district-level regressions of the change in the log share assigned names most associated with the culture to which we allocate the district based on surnames before 1800, against migration and industrialization. The independent variable in (1) is the change in the log number of people over 16 living outside the cluster divided by the number born in the district, 1851–1901. In (2), it is the change in the log number of people over 16 living in the district born outside the cluster, divided by the number born in the district, 1851–1901. Observations are weighted by the number allocated namescores in the 1851–1860 cohort. Robust standard errors in parentheses.

Table B.6: Relationship between migration and the popularity of the home culture

	Δ ln share S E culture			
	(1)	(2)	(3)	(4)
Δ ln Mf workers	0.062	0.038	0.013	−0.011
	(0.024)	(0.028)	(0.019)	(0.013)
Weights	1850s names	1900s names	1851 pop	none
N	809	809	809	809
R^2	0.048	0.024	0.001	0.001

This table presents the results of district-level regressions in which the independent variable is the change in log manufacturing employment between 1851 and 1901 and the dependent variable the change in the log share given names associated with the Southeast cluster between the 1851–1860 and 1901–1910 generations. In (1), as in Table 1, observations are weighed by the number allocated names in the 1851–1860 cohort, (2) weights by the number allocated names in the 1901–1910 cohort, (3) by 1851 population, and (4) is unweighted. Robust standard errors in parentheses.

Table B.7: Non-Robustness of Relationship Between Manufacturing Growth and Popularity of Southeast Culture

	Observed cultural change ($\Delta \ln \sigma$)					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted cultural change ($\Delta \ln \sigma$)						
from real wages (v_d^θ)	0.739	0.305				
	(0.021)	(0.054)				
from migration costs (δ_{od}^θ)			0.123	0.260		
			(0.036)	(0.027)		
from starting population (n_o)					0.388	0.445
					(0.026)	(0.102)
Cluster FE		x		x		x
N	7362	7362	7362	7362	7362	7362
R^2	0.234	0.521	0.004	0.530	0.043	0.523

This table shows OLS estimates at the district-by-cluster level. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination real wages v_d^θ in (1) and (2), migration costs in (3) and (4), and starting populations in (5) and (6), and the observed value for those born 1861–1895. The dependent variable is the change between the observed value for those born 1841–1860 and those born 1861–1895. Even-numbered models add cluster fixed effects. Standard errors clustered by district in parentheses.

Table B.8: Relationship between the change in log cultural choice shares, σ, relative to different 1851 counterfactuals

	Change in home culture ($\Delta \ln \sigma$)					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ ln real wage (v_d^θ)	−0.032	−0.056	−0.216	−0.161	−0.063	0.003
	(0.027)	(0.026)	(0.050)	(0.035)	(0.023)	(0.009)
Distance to London (100km)		−0.392	−0.867	−0.567	−0.354	−0.112
		(0.102)	(0.123)	(0.073)	(0.065)	(0.052)
$\Delta \ln v_d^\theta \times$ distance		0.078	0.183	0.120	0.081	0.014
		(0.025)	(0.028)	(0.016)	(0.015)	(0.012)
Counterfactual	Realized	Realized	All	Wages	Pop.	Migration costs
N	812	812	812	812	812	812
R^2	0.002	0.031	0.083	0.092	0.027	0.081

This table shows OLS estimates at the district-level of the relationship between wage growth and the popularity of the home culture across counterfactuals. The dependent variable is the log share choosing the home culture minus the log share predicted choosing the home culture in a specific counterfactual scenario. Models (1) and (2) examine the observed change between the 1841–1860 generation and the 1861–1895 generation. (3) examines the output from the counterfactual setting real wages, starting populations and migration costs to 1851 levels, (4) just real wages, (5) just populations, and (6) just migration costs. The independent variable is the change in real wages between 1851 and 1911, interacted with distance to London. All models include an intercept. Robust standard errors in parentheses.

Table B.9: Relationship between wage growth, distance to London, and counterfactual changes in the popularity of the home culture

	Change in log share migrating outside home culture			
	(1)	(2)	(3)	(4)
Δ ln real wage (v_d^θ)	−0.125	−0.296	0.035	0.135
	(0.049)	(0.021)	(0.024)	(0.018)
Distance to London (100km)	0.559	0.433	0.011	−0.127
	(0.230)	(0.065)	(0.116)	(0.064)
$\Delta \ln v_d^\theta \times$ distance	−0.100	−0.098	−0.000	0.048
	(0.052)	(0.015)	(0.026)	(0.015)
Counterfactual	All	Wages	Pop.	Migration costs
N	828	828	828	828
R^2	0.145	0.677	0.019	0.317

This table shows OLS estimates at the district-level of the relationship between wage growth and migration from the home culture across counterfactuals. The dependent variable is the log share migrating to locations outside the cluster to which we allocate the district based on historical surnames in the observed data, minus the predicted log share in the relevant counterfactual. (1) examines the output from the counterfactual setting real wages, starting populations and migration costs to 1851 levels, (2) just real wages, (3) just populations, and (64 just migration costs. The independent variable is the change in real wages between 1851 and 1911, interacted with distance to London. All models include an intercept. Robust standard errors in parentheses.

Table B.10: Relationship between wage growth, distance to London, and counterfactual changes in migration from the cluster

Table B.11: Data construction overview

	(1)	(2)	(3)
Task	Allocating districts to clusters	Calculating name scores	Calculating origin-destination-cluster migration flows
Data	Surnames of household heads born before 1800, allocated to district of birth	First names of people born 1841–1860, allocated to cluster of birth	People born 1861–1911, allocated to clusters by first name, to origin districts by birth, to destinations by residence
Source	1851 Census	1911 Census	1911 Census

Table B.12: Estimation overview

	(1)	(2)	(3)	(4)	(5)
Parameters	δ_{od}^{θ}, bundled $\left(\left(v_d \xi_d^k\right)^{\theta}\left(m_d^k\right)^{\alpha}\right)$	$\alpha, \left(\xi_d^k\right)^{\theta}, v_d^{\theta}$	Ω_o^k	φ, ψ_o^k	$v_{d,1851}^{\theta}, \delta_{od,1851}^{\theta}$
Estimation	Poisson regression	Two-stage least squares	Calculated directly from (1) and (2)	Poisson regression with control function	Poisson regression
Level	Origin-destination-culture	Destination-culture	Origin-culture	Origin-culture	Origin-destination
Data	1911 census	1911 census	1911 census	1911 census	1851 census

C EQUILIBRIUM UNIQUENESS

Allen, Arkolakis and Li (2020) provide conditions for equilibrium existence and uniqueness in spatial economic models. Their analysis focuses on an economic model with N agents and H types of interactions in which the equilibrium can be reduced to a set of $N \times H$ equations of the form

$$x_{ih} = \sum_{j=1}^{N} f_{ijh}(x_{j1}, ..., x_{jH}) \tag{7}$$

where x_{ih} is the equilibrium outcome for each agent in each interaction and f_{ijh} is a differentiable function that governs the interactions between agents. One can construct an $H \times H$ matrix of uniform bounds of the elasticities of these f_{ijh} functions,

$$\mathbf{A}_{hh'} = \sup_{i,j} \left(\left| \frac{\partial \ln f_{ijh}}{\partial \ln x_{jh'}} \right| \right)$$

Theorem 1 of Allen, Arkolakis and Li (2020) states that if the spectral radius of this matrix—the largest absolute value of its eigenvalues—is less than one, the equilibrium exists, is unique, and can be computed by iterating equation (7).

The main theorem requires that the f_{ijh} function only takes as arguments equilibrium outcomes pertaining to agent or location j and returns strictly positive values. Remark 1 gives an additional result for cases where f_{ijh} takes as arguments equilibrium outcomes pertaining to other agents or locations as well, and returns zeros. The results of Theorem 1 hold in such cases if one replaces the uniform bound on the elasticity with the uniform bound on the sum of elasticities.

To apply this theorem, we need to rewrite the equilibrium conditions of our model in a form consistent with equation (7).

Note that we can write (1), the number with culture k migrating from o to d, as

$$m_{od}^k = \frac{\left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha}{\Omega_o^k} \frac{\left(\Omega_o^k\right)^\varphi \psi_o^k}{\sum_{l=1}^K \left(\Omega_o^l\right)^\varphi \psi_o^l} n_o = \left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha n_o \frac{\left(\Omega_o^k\right)^{\varphi - 1} \psi_o^k}{\sum_{l=1}^K \left(\Omega_o^l\right)^\varphi \psi_o^l}$$

we can therefore write the total number of migrants of culture k in d as

$$m_d^k = \sum_{o=1}^{N} \left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha n_o \frac{\left(\Omega_o^k\right)^{\varphi - 1} \psi_o^k}{\sum_{l=1}^K \left(\Omega_o^l\right)^\varphi \psi_o^l}$$

Defining $q_d^k = \left(m_d^k\right)^{1-\alpha}$, we can rewrite this equation as

$$q_d^k = \sum_{o=1}^{N} \left(v_d \delta_{od} \xi_d^k\right)^{\theta} n_o \frac{\left(\Omega_o^k\right)^{\varphi-1} \psi_o^k}{\sum_{l=1}^{K} \left(\Omega_o^l\right)^{\varphi} \psi_o^l}$$

Similarly we can rewrite

$$\Omega_o^k = \sum_{j=1}^{N} \left(v_j \delta_{oj} \xi_j^k\right)^{\theta} \left(q_j^k\right)^{\frac{\alpha}{1-\alpha}}$$

We can thus express our equilibrium in terms of a set of $N \times K \times 2$ equations of the form $\Omega_o^k = \sum_{j=1}^{N} \sum_{l=1}^{K} g_{oj}^{k,l}(\mathbf{q}, \mathbf{\Omega})$ and $q_o^k = \sum_{j=1}^{N} \sum_{l=1}^{K} h_{oj}^{k,l}(\mathbf{q}, \mathbf{\Omega})$, where

$$g_{oj}^{k,l} = \left(v_j \delta_{oj} \xi_j^l\right)^{\theta} \left(q_j^l\right)^{\frac{\alpha}{1-\alpha}} \mathbf{1}_{\{l=k\}},\ h_{oj}^{k,l} = \left(v_o \delta_{jo} \xi_o^l\right)^{\theta} n_j \frac{\left(\Omega_j^l\right)^{\varphi-1} \psi_j^l}{\sum_{m=1}^{K} \left(\Omega_j^m\right)^{\varphi} \psi_j^l} \mathbf{1}_{\{l=k\}}.$$

In relation to Allen, Arkolakis and Li (2020)'s notation, here we treat $N \times K$ location-by-culture pairs as the set of locations and Ω_o^k and q_o^k as the two types of equilibrium outcomes. Note that the $g_{oj}^{k,l}$ function depends only on q_l^k, but can return zero values if $k \neq l$, and so is not consistent with equation (1) of Allen, Arkolakis and Li (2020). The $h_{oj}^{k,l}$ function depends on both Ω_j^l and Ω_j^m for other values $m \neq l$. We must apply Remark 1 of Allen, Arkolakis and Li (2020) which gives a condition related not to the uniform bound on the elasticity but to the uniform bound on the sum of the elasticities.

The relevant matrix of uniform bounds on the elasticities is then

$$\mathbf{A} = \begin{pmatrix} \sup_{o,k} \sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln \sum_{m=1}^{K} \sum_{d=1}^{N} h_{od}^{k,m}}{\partial \ln q_j^l} \right| & \sup_{o,k} \sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln \sum_{m=1}^{K} \sum_{d=1}^{N} h_{od}^{k,m}}{\partial \ln \Omega_j^l} \right| \\ \sup_{o,k} \sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln \sum_{m=1}^{K} \sum_{d=1}^{N} g_{od}^{k,m}}{\partial \ln q_j^l} \right| & \sup_{o,k} \sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln \sum_{m=1}^{K} \sum_{d=1}^{N} g_{od}^{k,m}}{\partial \ln \Omega_j^l} \right| \end{pmatrix}$$

For the top right-hand corner, note that $\sum_{m=1}^{K} \sum_{d=1}^{N} h_{od}^{k,m} = \sum_{d=1}^{N} h_{od}^{k,k} = q_o^k$

$$\frac{\partial \ln q_o^k}{\partial \ln \Omega_j^l} = \frac{\partial \ln q_o^k}{\partial \ln h_{oj}^{k,k}} \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l} = \frac{h_{oj}^{k,k}}{q_o^k} \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l}$$

Note,

$$\frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^k} = \varphi - 1 - \sigma_j^k \varphi,\ \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l} = -\sigma_j^l \varphi$$

where σ_j^l is the share choosing culture l in location j. The sum of these absolute values is then

$$\sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln q_o^k}{\partial \ln \Omega_j^l} \right| = \sum_{j=1}^{N} \frac{h_{oj}^{k,k}}{q_o^k} \left(\sum_{l=1}^{K} \left| \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l} \right| \right)$$

Writing out the term in parentheses

$$\sum_{l=1}^{K}\left|\frac{\partial\ln h_{oj}^{k,k}}{\partial\ln\Omega_j^l}\right| = \left|\varphi - 1 - \sigma_j^k\varphi\right| + \sum_{l\in\{1,\ldots,K\setminus k\}}\sigma_j^l\varphi$$

Given that $\sum_{l\in\{1,\ldots,K\setminus k\}}\sigma_j^l = 1-\sigma_j^k$, we can rewrite this equation as

$$\sum_{l=1}^{K}\left|\frac{\partial\ln h_{oj}^{k,k}}{\partial\ln\Omega_j^l}\right| = \left|\varphi\left(1-\sigma_j^k\right)-1\right| + \varphi\left(1-\sigma_j^k\right)$$

If $\varphi\left(1-\sigma_j^k\right)\leq 1$, this expression equals 1. If $\varphi\left(1-\sigma_j^k\right)>1$, it equals $2\varphi\left(1-\sigma_j^k\right)-1$, which is maximized at $\sigma_j^k=0$. This expression is therefore bounded by $\max\left(2\varphi-1,1\right)$ So we have

$$\sum_{j=1}^{N}\sum_{l=1}^{K}\left|\frac{\partial\ln q_o^k}{\partial\ln\Omega_j^l}\right| \leq \sum_{j=1}^{N}\frac{h_{oj}^{k,k}}{q_o^k}\max\left(2\varphi-1,1\right) = \max\left(2\varphi-1,1\right)$$

For bottom-left entry in $\mathbf{A}$,

$$\sum_{m=1}^{K}\sum_{d=1}^{N}g_{od}^{k,m} = \sum_{d=1}^{N}g_{od}^{k,k} = \Omega_o^k$$

$$\frac{\partial\ln\Omega_o^k}{\partial\ln q_j^k} = \frac{\partial\ln\Omega_o^k}{\partial\ln g_{od}^{k,k}}\frac{\partial\ln g_{od}^{k,k}}{\partial\ln q_j^k} = \frac{g_{od}^{k,k}}{\Omega_o^k}\frac{\alpha}{1-\alpha}$$

Note that $\frac{\partial\ln\Omega_o^k}{\partial\ln q_j^l}=0$ for $l\neq k$. Then

$$\sum_{j=1}^{N}\sum_{l=1}^{K}\left|\frac{\partial\ln\Omega_o^k}{\partial\ln q_j^l}\right| = \sum_{j=1}^{N}\left|\frac{\partial\ln\Omega_o^k}{\partial\ln q_j^k}\right| = \sum_{j=1}^{N}\frac{g_{od}^{k,k}}{\Omega_o^k}\frac{\alpha}{1-\alpha} = \frac{\alpha}{1-\alpha}$$

The two diagonals are zero because $\frac{\partial h_{od}^{k,m}}{\partial q_j^l}=0$ and $\frac{\partial g_{od}^{k,m}}{\partial\Omega_j^l}=0$ for all o,d,k,m,j,l.

We can therefore write

$$\mathbf{A} = \begin{pmatrix} 0 & \max(1,2\varphi-1) \\ \left|\frac{\alpha}{1-\alpha}\right| & 0 \end{pmatrix}$$

The spectral radius of this matrix is $\rho(\mathbf{A}) = \sqrt{\left|\frac{\alpha}{1-\alpha}\right|\max(1,2\varphi-1)}$. The condition that $\rho(\mathbf{A})<1$ will be satisfied if $\frac{\alpha}{1-\alpha}\max(1,2\varphi-1)<1$

D ADDITIONAL DETAIL ON INSTRUMENTAL VARIABLES

D.1 *Migration Flows in Table 2*

We want to regress the log share given names associated with a given culture in a location against the log share migrating from that location to districts historically associated with that culture. The endogeneity concern is that cultural choices and migration may both be influenced by unobserved preferences for the culture in question. We instrument for migration flows using coal deposits.

First, we aggregate migration data at the origin-destination level, and regress log migration flows against origin and destination fixed effects, log distance, and an indicator for the origin being the same as the destination:

$$\ln m_{od} = \gamma_d + \beta_1 \ln \text{distance}_{od} + \beta_2 \mathbf{1}_{\{o=d\}} + \gamma_o + \varepsilon_{od}$$

This gravity regression gives a set of migration cost coefficients, β_1, β_2, and destination fixed effects γ_d. We generate predicted destination fixed effects by regressing the estimated destination fixed effects against an indicator for coal presence:

$$\gamma_d = \delta_0 + \delta_1 \, \mathbf{1}_{\{d \text{ contains coal}\}} + \varepsilon_d$$

Given these various estimates, we can predict migration flows between each location:

$$\text{coal predicted } m_{od} = \frac{\exp\left(\hat{\delta}_0 + \hat{\delta}_1 \mathbf{1}_{\{d \text{ contains coal}\}} + \hat{\beta}_1 \ln \text{distance}_{od} + \hat{\beta}_2 \mathbf{1}_{\{o=d\}}\right)}{\sum_{j=1}^{N} \exp\left(\hat{\delta}_0 + \hat{\delta}_1 \mathbf{1}_{\{j \text{ contains coal}\}} + \hat{\beta}_1 \ln \text{distance}_{oj} + \hat{\beta}_2 \mathbf{1}_{\{o=j\}}\right)} n_o$$

where n_o is the number originating in o.

We then add up the number predicted to migrate to districts allocated to each culture, divide by the total number of migrants, and take the natural logarithm:

$$\ln \text{coal predicted } s_d^k = \ln \left(\frac{1}{n_o} \sum_d \text{coal predicted } m_{od} \mathbf{1}_{\{\text{historic culture}(d)=k\}} \right)$$

where $\mathbf{1}_{\{\text{historic culture}(d)=k\}}$ takes a value of 1 if the culture to which district d is allocated based on surnames in 1800 is k, and 0 if otherwise.

This gives the log share predicted to migrate to districts allocated to a given culture only due to coal deposits and proximity. To further isolate the component from coal deposits, we follow Borusyak and Hull (2020) and recenter the instrument. We randomly permute the vector of coal deposits 1,000 times. On each permutation, we calculate predicted migration flows and log migration shares. We then subtract the average of these permutate log migration shares from the instrument.

D.2 φ *Equation*

We construct the instrument as follows. First, we aggregate migration data at the origin-destination level, ignoring clusters, and regress log migration flows against origin and destination fixed effects, log distance, and an indicator for the origin being the same as the destination:

$$\ln m_{od} = \gamma_d + \beta_1 \ln \text{distance}_{od} + \beta_2 \mathbf{1}_{\{o=d\}} + \gamma_o + \varepsilon_{od}.$$

This gravity regression gives an alternative set of migration cost coefficients, β_1, β_2, and destination fixed effects γ_d. We generate predicted destination fixed effects by regressing the estimated destination fixed effects against an indicator for coal presence:

$$\gamma_d = \delta_0 + \delta_1 \, \mathbf{1}_{\{d \text{ contains coal}\}} + \varepsilon_d.$$

These estimates allow us to predict the migration pull to different locations, using just the presence of coal and geographic distance. We use these predicted migration incentives to construct on instrument for Ω_o^k:

$$\text{coal-predicted } \Omega_o^k = \sum_{d=1}^{N} \exp\left(\hat{\delta}_0 + \hat{\delta}_1 \mathbf{1}_{\{d \text{ contains coal}\}} + \hat{\beta}_1 \ln \text{distance}_{od} + \hat{\beta}_2 \mathbf{1}_{\{o=d\}}\right) \mathbf{1}_{\{\text{historic culture}(d)=k\}}$$

where $\hat{v}_d^\theta$ is the exponential of the coal-predicted destination utility, and the indicator function on the right takes a value of 1 if the district d is allocated to cluster k based on surnames before 1800. The logic is that coal and distance provide pull factors unrelated to cultural sorting, while locations allocated to a specific culture due to historic surnames will have higher cultural amenities ξ_d^k for that culture.

Again following Borusyak and Hull (2020) we recenter the instrument by permuting the vector of coal deposits 1,000 times, calculating Ω_o^k on each permutation using the permuted coal deposits, and subtracting the average $\ln \Omega_o^k$ over all permutations from our instrument. Recentering helps address the concern that districts close in space to a given historic culture may also have more similar cultural transmission tastes.

We then estimate φ with the control function as follows:

$$\ln \Omega_o^k = \beta \text{ recentered} \ln \text{coal-predicted } \Omega_o^k + \gamma_o + \gamma_{k(o)}^k + e_o^k$$

$$\ln \sigma_o^k = \varphi \ln \Omega_o^k + \delta \left(\ln \Omega_o^k - \ln \hat{\Omega}_o^k \right) + \gamma_o + \gamma_{k(o)}^k + \ln \psi_o^k$$

where $\gamma_{k(o)}^k$ is a fixed effect for the surname culture cluster o is located in interacted with the culture in question k, and $\left(\ln \Omega_o^k - \ln \hat{\Omega}_o^k \right)$ is the first-stage residual. The $\gamma_{k(o)}^k$ fixed effects help further address the concern that districts close to a given surname cluster might have stronger cultural transmission tastes for that cluster. We use our estimate of φ and observed σ_d^k and Ω_o^k to back out ψ_o^k.

D.3 α *Equation*

The coal-related influence on cultural choices that we use to instrument for Ω_o^k can also be used to instrument for destination populations and estimate α. We can predict migration flows from origins to destinations using coal, distance, and starting populations:

$$\text{coal-predicted } m_{od} = \frac{\exp\left(\hat{\delta}_0 + \hat{\delta}_1 \mathbf{1}_{\{d \text{ contains coal}\}} + \hat{\beta}_1 \ln \text{distance}_{od} + \hat{\beta}_2 \mathbf{1}_{\{o=d\}} \right)}{\sum_{j=1}^{N} \exp\left(\hat{\delta}_0 + \hat{\delta}_1 \mathbf{1}_{\{j \text{ contains coal}\}} + \hat{\beta}_1 \ln \text{distance}_{oj} + \hat{\beta}_2 \mathbf{1}_{\{o=j\}} \right)} n_o.$$

We then use these predicted migration flows—which do not take into account cultural differences—to predict how shocks to choices in the origins translate into destination populations:

$$\text{coal-predicted } m_d^k = \sum_{o=1}^{N} \text{coal-predicted } \Omega_o^k \cdot \text{coal-predicted } m_{od}.$$

This measure gives the weighted sum of shocks to cultural choices at the origin, weighted by the predicted number of migrants from that origin.

We also recenter this instrument following Borusyak and Hull (2020). As for the instrument for Ω_o^k, we permute the vector of coal deposits 1,000 times, calculate the instrument on each permutation, and substract the average of $\ln$ coal-predicted m_d^k from our instrument. When recentering the instrument, we only permute the coal deposits used to calculate coal-predicted Ω_o^k, not those used to calculate migration flows. Thus the average of the permuted instruments takes into account how coal infuences migration to specific origins, and so subtracting that average nets out such an effect. We then estimate the following equations by two-stage least squares:

$$\ln m_d^k = \beta \text{ recentered} \ln \text{coal-predicted } m_d^k + \gamma_d + \gamma_{k(d)}^k + e_d^k$$

$$\gamma_d^k = \alpha \ln \hat{m}_d^k + \gamma_d + \gamma_{k(d)}^k + \theta \ln \xi_d^k.$$

	ln religious distance		ln surname similarity		ln average distance	
	(1)	(2)	(3)	(4)	(5)	(6)
Recentered coal-predicted $\ln m_d^k$	-0.015		0.734		-0.200	
	(0.056)		(0.584)		(0.180)	
Recentered coal-predicted $\ln \Omega_o^k$		0.023		0.273		-0.115
		(0.044)		(0.418)		(0.107)
N	7425	7425	7264	7264	7452	7452
R^2	0.842	0.848	0.906	0.907	0.845	0.859

This table validates the exclusion restriction assumption for our instruments used to estimate α and φ by examining the relationship between the instruments and religious distance and surname similarity, measures of cultural proximity that correlate with ξ and ψ, the residuals in the second stage regressions, as well as average geographic distance to districts allocated to the culture in question. Data is at the district-cluster level. As in Table 4, observations are weighted by the relevant origin or destination population, and all models include district and cluster-by-home-cluster fixed effects. Standard errors clustered by district in parentheses.

Table D.1: Orthogonality of Instruments to Cultural Proximity

D.4 *Validating Instruments*

The exclusion restrictions necessary for our instrumental variables estimations are that the instrument for destination populations is uncorrelated with cultural amenities (ξ_d^k), and that the instrument for origin expected utilities is uncorrelated with cultural transmission tastes (ψ_o^k). While we cannot directly test that assumption, we can examine the relationship between our instruments and measures of cultural similarity that should be correlated with the taste and amenity terms. We examine the religious distance and surname similarity measures used in Section 6.6, and the average distance to districts allocated to the cluster in question, which should pick up other forms of cultural similarity that are geographically clustered.

Table D.1 regresses these measures of cultural similarity against our instruments, as in the first stages of our instrumental variables estimation routines. In all models, the coefficient on the instrument is close to zero and not statistically significant, which raises our confidence in the exclusion restriction.

E Bootstrapping elasticities

Table E.1 reports bootstrapped confidence intervals for the homophily (α) and culture (φ) elasticities. In our model, there are two sources of uncertainty. First, there uncertainty from the sampling of locations into our dataset. Second, origin-destination-culture migration flows are affected by transitory shocks that we treat as an error term in the estimation of distance elasticities and destination bundles of attributes. One can think of migration flows as a draw from a multinomial distribution. Randomness in the realization of this draw will affect our estimates.

To account for sampling uncertainty, one would normally use the nonparametric bootstrap, blocking by location. That is, one would resample locations or weights for locations to obtain a set of new datasets, and then run the entire estimation routine on each new dataset to obtain the sampling distribution of the quantity of interest. The key idea is that each resampled dataset approximates a draw from the same data generating process as the primary dataset. In our context that assumption would not be satisfied. Our model and counterfactual exercises indicate that changing the set of locations or populations alters individuals' cultural choices and migration decisions. If a bootstrap replication sampled London twice, we would expect more people to migrate there and the popularity of the southeast to increase. Observed cultural choice and migration shares that do not take into account this response would be inconsistent with the model.

Our preferred approach combines the clustered sampling aspect of the nonparametric bootstrap with the structure of our model. On each bootstrap replication we do the following:

1. Sample weights for origin locations from a Dirichlet distribution, as in a fractional random weight bootstrap

2. Multiply origin populations by these sampled weights

3. Solve for equilibrium migration flows and cultural choices given these new origin populations

4. Sample migration flows from a multinomial distribution, using the predicted share of members of culture k going from o to d as the choice probabilities

5. Estimate gravity model on these sampled migration flows to estimate destination fixed effects and migration cost parameters

6. Calculate instruments using these sampled migration flows

7. Use calculated instruments to estimate elasticities.

Doing so 1,000 times gives a distribution of estimated elasticities over resamples. We construct confidence intervals using the 2.5th and 97.5th percentiles of these distributions in Table E.1.

	$\ln m_d^k$	γ_d^k		$\ln \Omega_o^k$	$\ln \sigma_o^k$	
	(1)	(2)	(3)	(4)	(5)	(6)
Recentered coal-predicted $\ln m_d^k$	0.641					
	[0.095; 2.564]					
$\ln m_d^k$ [α]		0.604	0.521			
		[0.571; 0.634]	[0.385; 0.873]			
Recentered coal-predicted $\ln \Omega_o^k$				0.214		
				[0.070; 0.718]		
$\ln \Omega_o^k$ [φ]					2.185	1.803
					[2.038; 2.390]	[0.844; 2.355]
First-stage residuals						0.388
						[−0.040; 3.510]
Model	First stage	OLS	TSLS	First stage	PPML	Control function
N	7418	7418	7418	7452	7452	7452
R^2	0.989	0.999	0.999	0.998		

This table shows estimates of α and φ with confidence intervals from a semiparametric bootstrap. All models are estimated at the district-by-culture level. Models (1) and (4) show the first stage, (2) and (5) the non-instrumented second stage and (3) and (6) the instrumented second stage. (1), (2), and (4) are estimated by OLS, (3) by TSLS, and (5) and (6) by Poisson Pseudo-Maximum Likelihood, with (6) implementing the control function approach by including the residuals from (4). In models (2) and (3) the dependent variable is the destination-by-culture fixed effect from an origin-by-destination-by-culture gravity model, in (5) and (6) it is the log share of people born in the district assigned to the culture. All models include fixed effects for the district and the culture interacted with the historic culture of the district. Standard errors are calculated through a semiparametric bootstrap clustered by location. On each bootstrap iteration we sample weights for origin locations, solve for equilibrium migration and cultural choices, draw origin-destination-culture migration flows from a multinomial distribution given these choices and run the entire estimation routine on the resulting dataset. Confidence intervals are the 2.5th and 97.5th percentiles of the bootstrap distribution.

Table E.1: Estimates of model elasticities with bootstrapped confidence intervals

F Estimating 1851 economic fundamentals

This appendix details how we estimate real wages and migration costs using 1851 data. In the main estimation, we use origin-destination-culture migration flows to separately estimate the real wage v_d^θ and cultural amenities ξ_d^k. For 1851, we do not have data on names and so cannot allocate individuals to cultures. We estimate the following gravity model using origin-destination migration flows from 1851:

$$\ln m_{od,1851} = \gamma_{d,1851} + \beta_{1,1851} \ln \text{distance}_{od} + \beta_{2,1851} \mathbf{1}_{\{o=d\}} + \gamma_{o,1851} + \varepsilon_{od,1851} \tag{8}$$

The parameters $\beta_{1,1851}$ and $\beta_{2,1851}$ pin down migration costs as a function of geographic distance and an indicator that the origin is the same as the destination, much as in our main estimation in Section 6.1.

In a model without cultural differentiation, the destination fixed effect $\gamma_{d,1851}$ would correspond to $\theta \ln v_{d,1851}$, the real wage. However, in our model agents have a preference for being in a location with other members of the same culture, so some of the migration pull to a particular location is due to the number of others migrating there. To estimate the real wage, we need to rescale these destination fixed effects. We estimate the equivalent of (8) using 1911 data, and then regress the log real wages, $\theta \ln v_d$, estimated for 1911 in Section 6.2 against $\gamma_{d,1911}$, the destination fixed effects from this 1911 gravity model. In that regression, the coefficient on the destination fixed effects is 0.43 and the R^2 is 0.88. We then rescale the destination fixed effects from the 1851 model by 0.43 to account for this sorting effect: $v_{d,1851}^\theta = \exp(\gamma_{d,1851})^{0.43}$

G model implied elasticities

Table G.1 examines whether model-implied elasticities moderate the relationship between changes in real wages and cultural choices. The model-implied elasticities in question, described in more detail in Section 8 are the predicted change in the popularity of a given culture in a given district in a simulation that increases the district's real wage by 1%. We take the log of this predicted change and divide by ln(1.01) so it can be interpreted as an elasticity. These elasticities should tell us whether and how much a culture grows in a district in response to economic change in the district. We test this claim by regressing the change in log cultural choices for each culture and district against the change in log real wages in that district, which we interact with the model-implied elasticity. The coefficient of interest is on the interaction between changes in real wages and the elasticity. A positive coefficient would indicate that increasing real wages leads to increases in the prevalence of the cultures with positive culture-wage elasticities. In models (1)–(3) we find this to be the case. In (5) we find that this effect also holds when subsetting to the home culture, the culture to which we allocate the district based on historical surnames. In (4) we find that wage growth in districts further from London preserved the home culture, in (6) we find that this relationship is driven by greater home-culture wage elasticities further from London.

	Change in cultural choice ($\Delta \ln \sigma$) $\times$ 100					
	(1)	(2)	(3)	(4)	(5)	(6)
Model-implied elasticity	-2.122	-1.740	-1.533		-8.044	-8.022
	(0.813)	(0.743)	(0.635)		(2.581)	(2.654)
Distance to London (km)				-0.392		-0.236
				(0.102)		(0.098)
Δ ln real wage ($\Delta \theta \ln v_d$)	2.091	1.968		-5.574	-3.830	-2.572
	(0.692)	(0.698)		(2.575)	(2.343)	(2.065)
$\Delta \theta \ln v_d \times$ elasticity	0.427	0.330	0.309		1.867	1.912
	(0.189)	(0.176)	(0.151)		(0.639)	(0.656)
$\Delta \theta \ln v_d \times$ distance to London				0.078		0.039
				(0.025)		(0.023)
Clusters	All	All	All	Home	Home	Home
Cluster FE		x	x			
District FE			x			
N	7362	7362	7362	812	812	812
R^2	0.006	0.522	0.586	0.031	0.029	0.058

This table shows OLS estimates at the district-by-cluster level. The dependent variable is the change in the log share choosing the culture in the district, between the 1841–1860 cohort and the 1861–1895 cohort, multiplied by 100 to aid legibility. Models (1)–(3) are estimated using all clusters and interact changes in real wages with the model-implied elasticity of σ to v_d^θ, estimated by simulating a 1% increase in v_d^θ separately for each district. The coefficient of interest is this interaction, which should be positive. Models (2) and (3) include cluster fixed effects, (3) includes district fixed effects which are collinear with the change in real wages. Models (4)–(6) subset to the culture cluster the district is allocated based on historical surnames. (4) shows that further from London, wage growth was more positively associated with the preservation of the home culture, (5) shows the interaction between wage growth and the model-implied elasticity, (6) includes both and shows that the interaction effect of being far from London is attributable to the model-implied elasticity. In (1) and (4)–(6) an intercept is not shown. Standard errors clustered by district in (1)–(3) or robust standard errors in (4)–(6) in parentheses.

Table G.1: Relationship between industrialization and cultural change, moderated by model-implied elasticities

	General equilibrium elasticity			
	(1)	(2)	(3)	(4)
Partial equilibrium elasticity	5.650	344.083	5.523	344.234
	(0.763)	(5.424)	(0.755)	(5.307)
District FE		x		x
Cluster FE			x	x
N	7345	7345	7345	7345
R^2	0.011	0.924	0.018	0.925

This table shows OLS estimates at the district-by-cluster level. The dependent variable is the model-implied general equilibrium elasticity of σ_o^k to v_o^θ, the independent variable is the partial equilibrium elasticity, calculated as the gravity-model predicted share of people of culture k in o not migrating from location o. Models (2) and (4) include district fixed effects, (3) and (4) cluster fixed effects. Standard errors clustered by district in parentheses.

Table G.2: Relationship between partial and general equilibrium elasticities

H ADDITIONAL ANALYSES

H.1 *Effects of Cultural Homogeneity on Migration and Economic Outcomes*

This section examines whether rising cultural homogeneity affected economic outcomes. Because culture-specific preferences and homophily influenced where people migrated, changing the popularity of different cultures should alter the spatial distribution of population, and, if doing so directs more people towards locations with higher real wages, increase average income (v_d). Table H.1 shows the results of different counterfactual simulations, primarily changing the ψ_o^k matrix of place-specific preferences for choosing particular cultures. Total assimilation into the Southeast culture, achieved by setting all non-Southeast ψ_o^k values to zero, increases the average v_d^θ by 0.76%.

If we think of v_d as representing the real wage in location d, under constant returns to scale, then given an estimate of θ, the elasticity of migration to wages, we can infer the effect on average wages. Tombe and Zhu (2019), using data from China, estimate elasticities of migration to wages between 1.2 and 1.6, and use 1.5 in their analysis, which would suggest that the estimate in the first row of Table H.1 corresponds to around a 0.5% increase in average wages. Morten and Oliveira (2023) estimate a migration elasticity of 4.5 in Brazil, corresponding to a 0.2% increase in average wages in response to full homogenization into the Southeast culture. Caliendo et al. (2021) estimate a much lower elasticity of 0.5—corresponding to a 1.5% increase in average wages—albeit in a dynamic model of migration in Europe. The positive effect of homogenization into the Southeast culture on economic output is due to the correlation between cultural amenities (ξ_d^k) specific to the Southeast culture and real wages: increasing the prevalence of the Southeast culture directs more people to high real wage locations. If we set all culture-specific destination migration preferences to 1, we find that cultural homogeneity in fact decreased average v_d^θ. Figure H.1 shows that both counterfactuals redistribute population across regions, away from rural parts of the North and Wales. More moderate changes to the popularity of the Southeast cluster have more moderate effects, while rescaling ψ_o^k so that the average value for all cultures is equal decreases average v_d^θ.

These small effects of homogenization on economic output make sense given the paper's other findings. Distance-related migration costs that help preserve peripheral cultures by diminishing out-migration also keep workers in low-productivity peripheral regions, even in the absence of cultural sorting and preferences. The model already assumes that individuals migrate towards higher-wage locations, which limits the extent to which eliminating other factors can shift migration towards such locations.

Table H.1: Counterfactual estimates from homogenizing culture

Counterfactual	%Δ Ave. v_d^θ	%Δ share migrant
Removing non-S E ψ_o^k	0.757	0.702
Removing non-S E ψ_o^k and fixing $\xi_d^k = 1$	-0.152	0.249
Doubling S E ψ_o^k	0.755	0.696
Halving S E ψ_o^k	-0.317	-0.015
Fixing ψ_o^k to have same average for all cultures	-0.116	0.144

This table shows percentage changes in the average v_d^θ and share migrating under different counterfactual scenarios. The first replaces all cultural choice preferences ψ_o^k with zero for cultures other than the Southeast. The second adds to this specification by also fixing ξ_d^k to one for all destinations. The third replaces ψ_o^k with double its value for the Southeast culture, the fourth with half its value. The fifth rescales ψ_o^k so that the average value for all cultures is equal.

H.2 *Measuring Transportation Costs*

In the main version of the model, transportation costs are simply a function of geographic distance and an indicator that the origin is the same as the destination. In this section we explore directly calculating transportation costs using shapefiles for railways, stations, shipping routes, and ports.

We use data on railways in 1851 and 1881 from Satchell et al. (2023) and Marti-Henneberg et al. (2023), stations in 1851 and 1881 from Marti-Henneberg et al. (2018) and Henneberg et al. (2018), and ports and shipping routes from Alvarez-Palau et al. (2019).

Our method closely follows Donaldson and Hornbeck (2016) in constructing a matrix of transportation costs between Registration District Centroids. With this data for 1851 and 1881, we create a network database in which nodes are locations—district centroids, stations, and ports—and edges are railways, shipping routes, and land routes connecting them. In addition to railways and shipping we allow any two nodes within 100km of each other to be connected by land. We weight these edges using the transportation cost parameters from Donaldson and Hornbeck (2016) and then use Dijkstra's algorithm to calculate the least cost path between any two district centroids. One should note that Donaldson and Hornbeck are studying trade costs in the US during the same period, and so it is possible that the relative cost of different modes of transport for shipping goods differs from the relative cost for transporting people.

Figure H.2 plots these calculated transport costs against distance in 1851 and 1881. As expected, transportation costs strongly correlate with geographic distance. The expansion of the railway network served to reduce transportation costs at intermediate-level distances.

Table H.2 shows the relationship between these transportation costs and migration flows. These models regress log migration flows against log transportation costs, with origin-year and destination-year fixed effects. In the context of the gravity model of migration, one

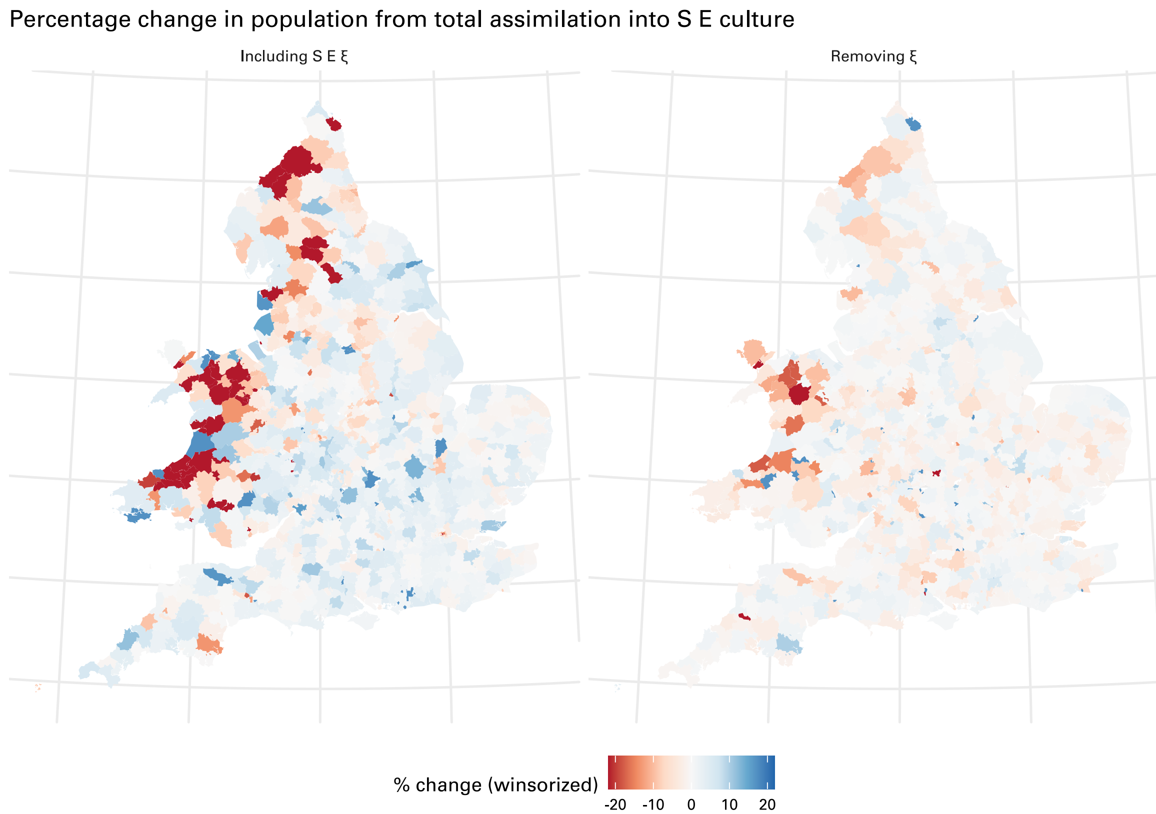

Figure H.1: Changes in population distribution from homogenizing culture

These maps show the percentage change in destination populations under counterfactual simulations that replace cultural preference terms ψ_o^k for all cultures except the Southeast culture with zero. The figure on the right also replaces all culture-specific migration preferences $\left(\xi_d^k\right)^\theta$ with 1.

can think of the resulting estimates as indicating the relationship between transportation costs and the migration costs implied by migration flows. In model (1) we find a strong negative relationship between transportation costs and migration flows, consistent with the idea that transportation costs influenced migration costs. However, in (2) we find that controlling for geographic distance shrink the coefficient on transportation costs to zero. The implication is that (1) picks up the effects of distance and not a distinct effect of infrastructure. In (3) we add origin-destination fixed effects and similarly find a null effect of transportation costs on migration flows. Note that in all these specifications, we would expect the coefficient on transportation costs to be biased away from zero in the negative direction due to the endogeneity of infrastructure construction: demand for migration from an origin to a destination could drive investment in connections between the two locations.

Our conclusion from this exercise is that geographic distance provides a better measure of migration costs than measures of transportation costs. In addition to possible error in the measurement of transportation costs, it is plausible that distance captures something related to ease of access or familiarity that is distinct from the physical cost of transportation.

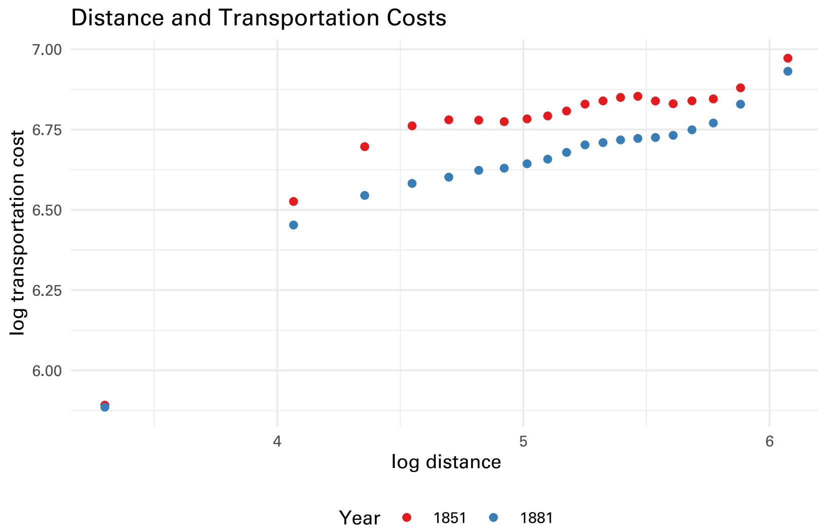

Figure H.2: Distance and Transport Costs, 1851 and 1881

Binned scatterplot of log transportation costs (y axis) in 1851 and 1881 against log distance.

H.3 *Relationship Between Amenities and Wages*

Do cultural amenities capture just taste, or also local wage differentials? If members of a given culture are paid more than members of other cultures in a given location, that would induce members of that culture in particular to migrate there, and would be picked up by the cultural amenity parameter in our model. Such a wage difference could emerge if there is demand for specific skills in a location, which members of some cultures but not others tend to have, or if there is taste-based discrimination against members of certain cultures in the labor market.

Table H.3 examines this claim, using HISCAM scores (Lambert et al., 2013). These provide a measure of the socioeconomic status of each occupation. Our dependent variable is the average HISCAM score, at the culture-district level, for men born 1861–1895. A higher value implies that men from that culture living in that location tended to have higher status occupation. Our data does not allow us to examine whether men were paid more to do the same job. We regress average HISCAM scores against real wages (v_d^θ), cultural amenities (ξ_d^k), and transmission tastes (ψ_o^k), weighting by the number of individuals in each cell. This specification gives equivalent coefficients to an individual-level regression. We find an ambiguous relationship between cultural amenities and transmission tastes, and HISCAM

	log migrants		
	(1)	(2)	(3)
log transportation cost	−2.422	0.105	−0.089
	(0.098)	(0.081)	(0.060)
log distance		−1.743	
		(0.071)	
Origin-year FE	x	x	x
Destination-year FE	x	x	x
Origin-destination FE			x
N	1382742	1382742	795729

This table presents evidence of the relationship between transportation cost, distance, and migration flows in the 1851 and 1881 censuses. Models are estimated by Poisson Pseudo-Maximum Likelihood. Data is at the origin-destination-year level. The dependent variable is the log number migrating from the origin to the destination in that year's census. The independent variable of interest is the log transportation cost. All models include origin-year and destination-year fixed effects. Model (2) controls for the log distance between the origin and destination. (3) adds origin-destination fixed effects, which are collinear with distance. Standard errors clustered by origin and destination in parentheses.

Table H.2: Null relationship between transport costs and migration

scores: in two specifications the taste parameter is positively correlated with status, in two they are negatively correlated, though only the negative coefficients are statistically significant. These results provides evidence against the idea that cultural amenities capture positive local wage differentials.

A long tradition in urban economics dating back to Roback (1982) argues that amenities that affect workers' preferences but not firm productivity should lead to lower wages. The logic is that there is a compensating differential: if in equilibrium the marginal worker is indifferent across locations, a factor that makes a location more desirable will attract more workers and push down wages. While we do not explicitly model a culture-by-location process that would generate such a compensating differential, Bryan and Morten (2019) suggest one way to reconcile a negative relationship between cultural amenities and wages with our theoretical and empirical framework. In their model of migration, the individual-by-destination Frèchet shock (the equivalent of ε_{id} in our model) is interpreted as a location-specific skill. A higher amenity value for members of a given culture in a given location serves to attract members of that culture with lower Frèchet shocks for that location, who have less-valuable location-specific skills and earn less.

	Average economic status			
	(1)	(2)	(3)	(4)
log real wage (v_d^θ)	0.312		0.283	
	(0.364)		(0.368)	
log cultural amenity (ξ)	-5.109	0.332		
	(1.285)	(0.303)		
log cultural transmission taste (ψ)			0.733	-0.809
			(0.439)	(0.122)
Cluster FE	x	x	x	x
District FE		x		x
N	7385	7385	7375	7375
R^2	0.080	0.979	0.071	0.981

This table shows OLS estimates at the district-cluster level. The dependent variable is the average economic status inferred using the HISCAM score corresponding to each occupation, for men born 1861–1895 resident in the district, allocated to clusters by their first names. In models (1) and (2) the independent variable is the destination (log) cultural amenity (ξ), in (3) and (4) the (log) cultural transmission taste (ψ). All models include cluster fixed effects. Models (1) and (3) also include the destination (log) real wage (v_d^θ), (2) and (4) add district fixed effects which are collinear with real wages. Observations are weighted by the number of individuals for which occupational status data is available Standard errors clustered by district in parentheses.

Table H.3: Relationship between cultural amenities, real wages, transmission tastes, and economic status

H.4 *Accounting for Emigration*

During this period, vast numbers of English and Welsh people emigrated, especially to the US, Canada, Australia, and New Zealand. These emigrants are not in our dataset, and to the best of our knowledge, there is no comprehensive data on emigrants. The question is how emigration affects our results. Here we show that our theory and evidence related to internal migration is unaffected by the existence of emigration. The potential for emigration does affect cultural choices, but in our framework it is captured by the location-by-culture cultural transmission tastes. Our counterfactual exercises must be interpreted as holding the ratio of domestic to international migration opportunities fixed. This assumption is appropriate given we are interested in examining how changing patterns of industrialization and internal migration affect identity choice, not in how changing international opportunities affect cultural choices.

Suppose that in addition to the locations we observe, there are a number of unobserved locations, to which people can also migrate. In the main body of the paper, locations 1 through N are domestic. Let us add $F - N$ foreign locations, so that there are F total locations, where 1 through N are domestic and $N + 1$ through F are foreign.

Denote the number of people born in o assigned to culture k, including those who emigrate, with $n_o^k + e_o^k$, where e_o^k refers to the number of emigrants from o of culture k, and n_o^k, as in our analysis, is the number originating in o of culture k who remain in England and Wales. Then the number migrating to each location d from domestic locations (Equation (1)) becomes

$$m_{od}^k = \frac{\left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha}{\sum_{f=1}^{F} \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha} \left(n_o^k + e_o^k\right). \tag{9}$$

Note that the number from o of culture k who remain in England and Wales, n_o^k, is the sum over domestic locations:

$$n_o^k = \sum_{j=1}^{N} m_{oj}^k = \frac{\sum_{j=1}^{N} \left(v_j \delta_{oj} \xi_j^k\right)^\theta \left(m_j^k\right)^\alpha}{\sum_{f=1}^{F} \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha} \left(n_o^k + e_o^k\right),$$

which we rearrange as follows

$$\frac{n_o^k}{n_o^k + e_o^k} \frac{\sum_{f=1}^{F} \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha}{\sum_{j=1}^{N} \left(v_j \delta_{oj} \xi_j^k\right)^\theta \left(m_j^k\right)^\alpha} = 1.$$

Multiplying (9) by the left-hand side of this equation gives

$$m_{od}^{k}=\underbrace{\frac{\left(v_d\delta_{od}\xi_d^k\right)^{\theta}\left(m_d^k\right)^{\alpha}}{\sum_{j=1}^{N}\left(v_j\delta_{oj}\xi_j^k\right)^{\theta}\left(m_j^k\right)^{\alpha}}n_o^k}_{m_{od}^{k}\text{ in (1)}}\underbrace{\frac{\sum_{f=1}^{F}\left(v_f\delta_{of}\xi_f^k\right)^{\theta}\left(m_f^k\right)^{\alpha}}{\sum_{f=1}^{F}\left(v_f\delta_{of}\xi_f^k\right)^{\theta}\left(m_f^k\right)^{\alpha}}\frac{n_o^k+e_o^k}{n_o^k+e_o^k}}_{=1}.$$

This exercise shows that allowing for emigration and unobserved destination locations does not affect the part of our model that analyzes migration choices and so does not affect our calculations of the relevant v_d, δ_{od} and ξ_d^k components. The logic is that we are already implicitly conditioning on people remaining in England and Wales.

Counterfactuals that set wages back to 1850s levels should affect emigration, for these analyses one must in effect assume that we are holding the ratio of domestic to foreign migration opportunities fixed.

Including foreign locations does alter the expressions for cultural choices. If foreign migration opportunities also matter for cultural choices, then in addition to the domestic Ω_o^k, there is an additional foreign term:

$$\Omega_o^{k,t}:=\sum_{f=1}^{F}\left(v_f\delta_{of}\xi_f^k\right)^{\theta}\left(m_f^k\right)^{\alpha}=\Omega_o^k+\Omega_o^{k,f}.$$

Here $\Omega_o^{k,t}$ is the culture-specific expected utility inclusive of emigration, and

$$\Omega_o^{k,f}:=\sum_{f=N+1}^{F}\left(v_f\delta_{of}\xi_f^k\right)^{\theta}\left(m_f^k\right)^{\alpha}$$

is the expected utility just from foreign locations. Modifying the expression for cultural choices in (3) to include foreign migration opportunities gives

$$\sigma_o^k=\frac{\left(\Omega_o^{k,t}\right)^{\varphi}\psi_o^k}{\sum_{l=1}^{K}\left(\Omega_o^{l,t}\right)^{\varphi}\psi_l^k}.$$

Defining the ratio of total to domestic culture-specific expected utility as $\omega_o^k:=\frac{\Omega_o^{k,t}}{\Omega_o^k}$, we have

$$\sigma_o^k=\frac{\left(\Omega_o^k\right)^{\varphi}\left(\omega_o^k\right)^{\varphi}\psi_o^k}{\sum_{l=1}^{K}\left(\Omega_o^l\right)^{\varphi}\left(\omega_o^l\right)^{\varphi}\psi_l^k}.$$

Taking logarithms as in (5) to create an estimating equation, we see that the error term in this regression now includes the ω_o^k term:

$$\ln \sigma_o^k = \varphi \ln \Omega_o^k - \underbrace{\ln \left(\sum_{l=1}^{K} \left(\Omega_o^l\right)^{\varphi} \left(\omega_o^l\right)^{\varphi} \psi_l^k \right)}_{\text{place FE}} + \underbrace{\varphi \ln \omega_o^k + \ln \psi_o^k}_{\text{Error}}.$$

The implication is that our estimates of the cultural transmission taste ψ_o^k also include the ratio of total to domestic migration opportunities, ω_o^k. If it were easier to practice one culture in the United States than others, then locations with greater ease of migration to the United States would appear to have a stronger cultural transmission taste for that culture. Our counterfactuals should then be interpreted as providing the estimated effects of changes in the domestic economy, holding fixed the ratio of domestic to foreign opportunities.

H.5 *Alternative Elasticities*

A key result in this paper is that the change in cultural choices going from a counterfactual estimated using 1851 economic fundamentals to the observed 1911 data matches specific observed cultural changes. We show that the culture of the Southeast of England expands, especially in places further from London, while peripheral home cultures decline, an effect moderated by the presence of coal. Here we investigate whether the ability of our model to reproduce these patterns is sensitive to the estimated homophily (α) and culture (φ) elasticities.

For an arbitrary pair of elasticities, we back out the other parameters of the model, and solve for counterfactual cultural choices under 1851 real wages, populations, and migration costs. Figure H.3 plots the equivalent of Figure 10A at each of these pairs of elasticities, Figure H.4 plots the equivalent of Figure 11A. Note that in both figures the scale of the y axis varies across rows. We only include pairs of elasticities such that $\alpha\varphi < 1$, because for values above that bound the equilibrium is not unique and so simulation results are hard to interpret. The broad takeaway is that the qualitative predictions of an increase in the Southeast culture further from London, and a decline in the home culture further from London moderated by the presence of coal hold across non-zero values of these elasticities. However, the magnitude of the predicted changes, and the steepness of the slope linking distance to London and predicted changes, increases considerably as we increase the magnitude of the elasticities.

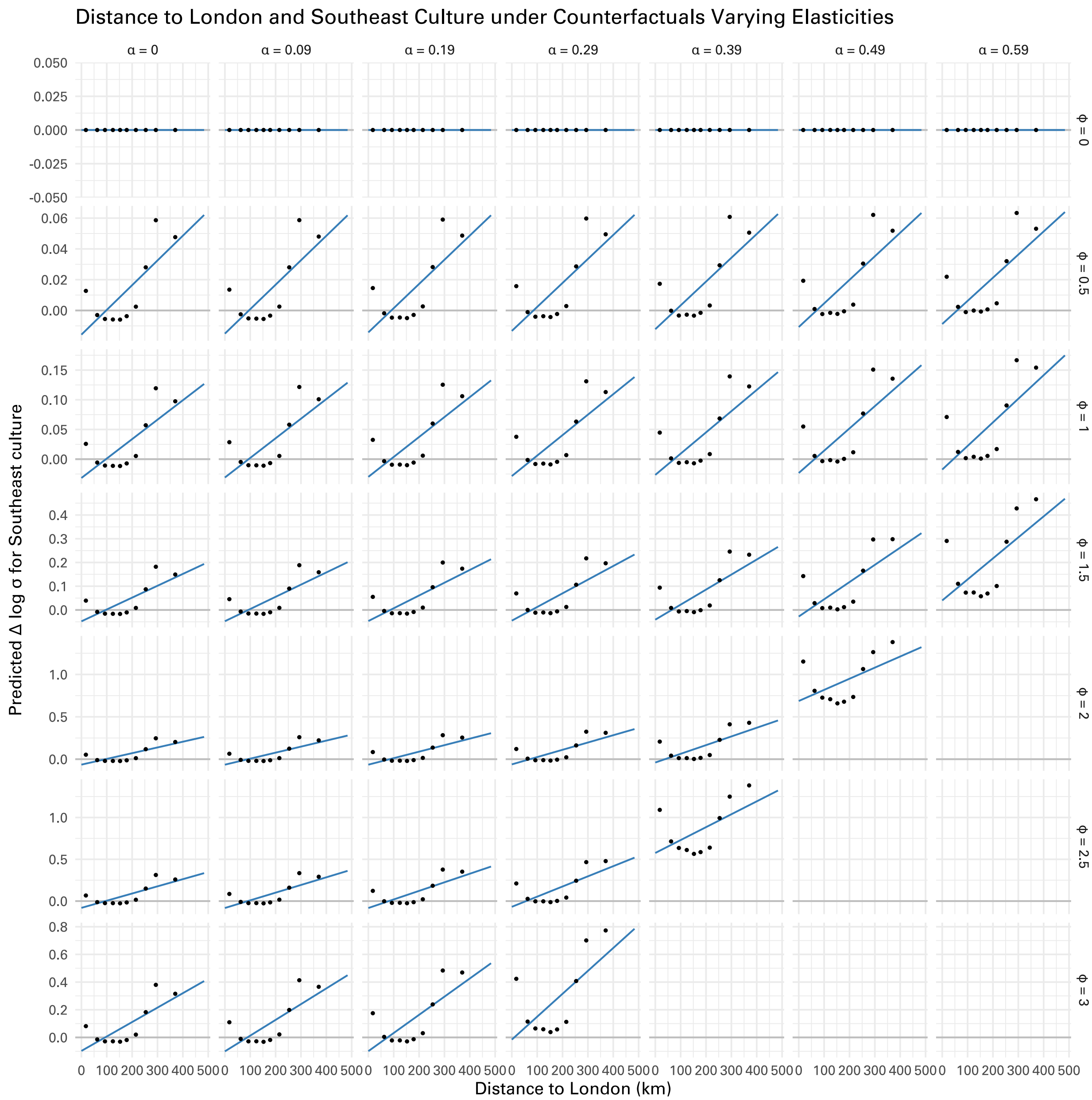

Figure H.3: Relationship between distance to London and the rise of the Southeast culture under different elasticities

Each grid cell shows the binned scatter relationship between distance to London and the log share choosing the Southeast culture in the observed data minus that in a counterfactual estimated using 1851 real wages, populations and migration costs. Each cell varies the homophily (α) and culture elasticities (φ). Instead of estimating these elasticities using instrumental variables, we fix them at a given value and proceed to back out the other parameters of the data, then solve for counterfactuals using these assumed elasticities and backed out parameters. Columns vary α, rows φ. Note that the scale of the y axis varies across rows. We do not calculate counterfactuals for values of $\alpha\varphi \geq 1$ as the equilibrium is non-unique and unstable.

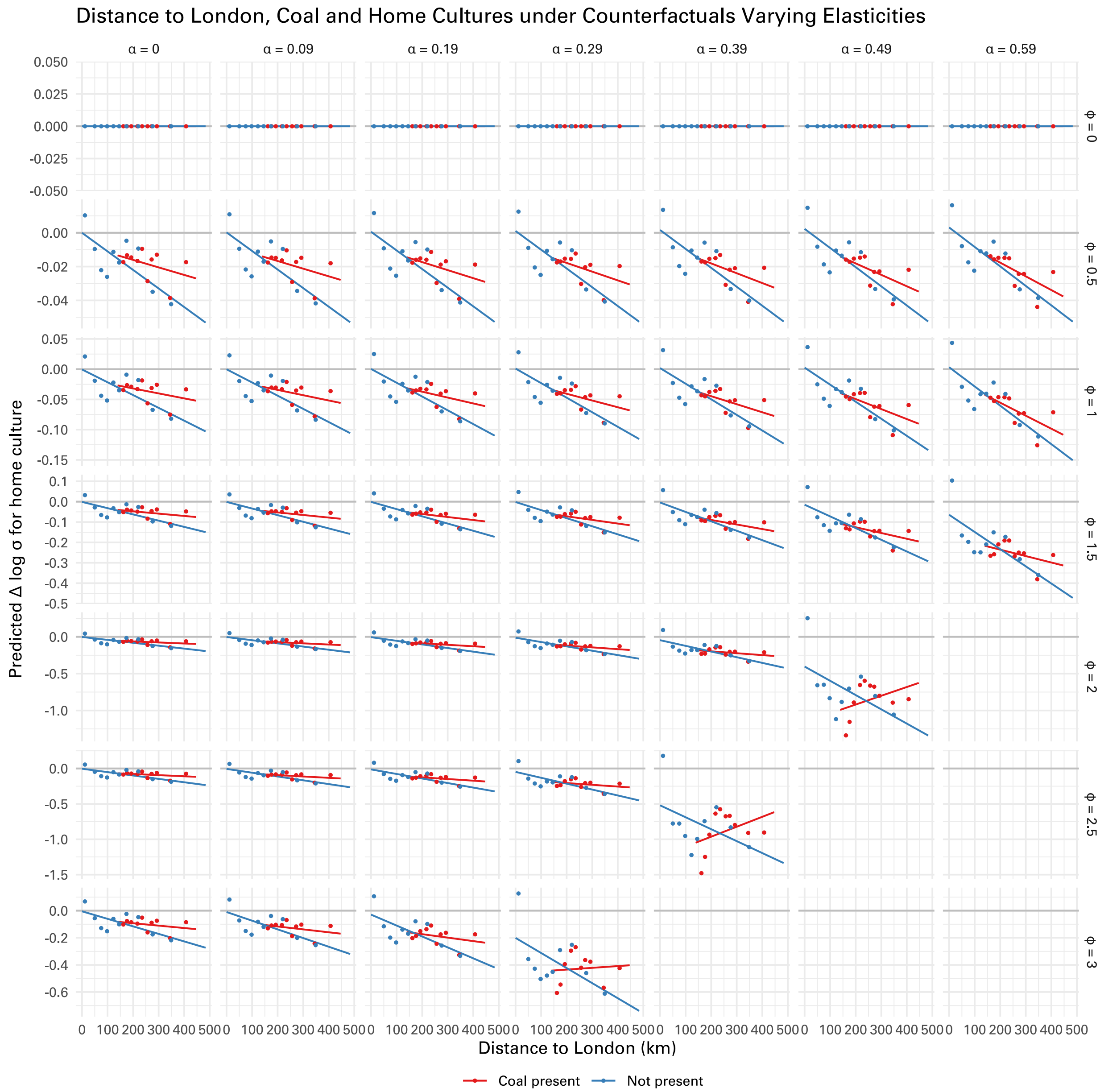

Figure H.4: Relationship between distance to London, coal, and the decline of the home culture under different elasticities

Each grid cell shows the binned scatter relationship between distance to London and the log share choosing the home culture in the observed data minus that in a counterfactual estimated using 1851 real wages, populations and migration costs, subset by whether the district contains coal. Each cell varies the homophily (α) and culture elasticities (φ). Instead of estimating these elasticities using instrumental variables, we fix them at a given value and proceed to back out the other parameters of the data, then solve for counterfactuals using these assumed elasticities and backed out parameters. Columns vary α, rows φ. Note that the scale of the y axis varies across rows. We do not calculate counterfactuals for values of $\alpha\varphi \geq 1$ as the equilibrium is non-unique and unstable.

H.6 *Alternative Cultural Clusters*

In this section we examine how results from the paper change using alternative cultural clusters. We focus on two sets of clusters. First, we group counties into the regions used to coordinate civil defence during the Second World War. This grouping was made after the changes we study, but it was one of the first administrative geographies above the county level, and gives a similar number of regions to the cultural clusters we estimate. The point is not that this grouping combines culturally-similar regions, but rather provides a way of combining adjacent counties. The left panel of Figure H.5 shows these regions.

Second, we estimate clusters using data on medieval English from *A Linguistic Atlas of Late Medieval English* (Benskin et al., 2013). This source codes the appearance of linguistic features in geocoded medieval texts. We create a measure of the similarity between these sources based on the share of common linguistic features, and then run the Louvain graphical clustering algorithm on the similarity matrix. Relative to the spectral algorithm we use with surname data, the Louvain algorithm is well-suited to data that already encodes the distance between entries. It also automatically estimates the optimal number of clusters. We allocate districts to clusters based on the most common cluster among the 7 sources closest to the district. Because the data only applies to England, we allocate Wales to its own cluster. The right panel of Figure H.5 shows these clusters.

The pattern we observe in Figure 4, whereby by the start of the twentieth century, most of England was sucked into the Southeast-English culture, with exceptions in Wales and the Northeast, holds under these different clusters. Figures H.6 and H.7 both show the cluster with the highest name score among those born in each district 1851–1860 and 1901–1910, calculated using the two alternative clusters. Using civil defence regions, the hegemonic culture by 1910 is London, with linguistic cultures, the large Southern English cluster. These patterns suggest that the rise of the Southeast that we observe in the main text is primarily the rise of London.

We examine how both our reduced-form and structural findings change as we change the cultural clusters. For the reduced-form analyses in Table 2, we simply re-calculate name scores, migration flows from districts to clusters, and our coal-based instrument, using the different cluster allocations. For the structural analyses, we calculate name scores and cluster-by-migration flows using the new clusters, and calibrate the model using the elasticities we estimate in the main body of the paper. We then examine whether the cultural changes predicted by the model conform to actual cultural changes, as in Table 5. We re-run the main counterfactual simulations and examine the new predictions for the aggregate changes of the hegemonic culture (London in the case of the civil defence regions, and the large Southern cluster in the case of the linguistic clusters).

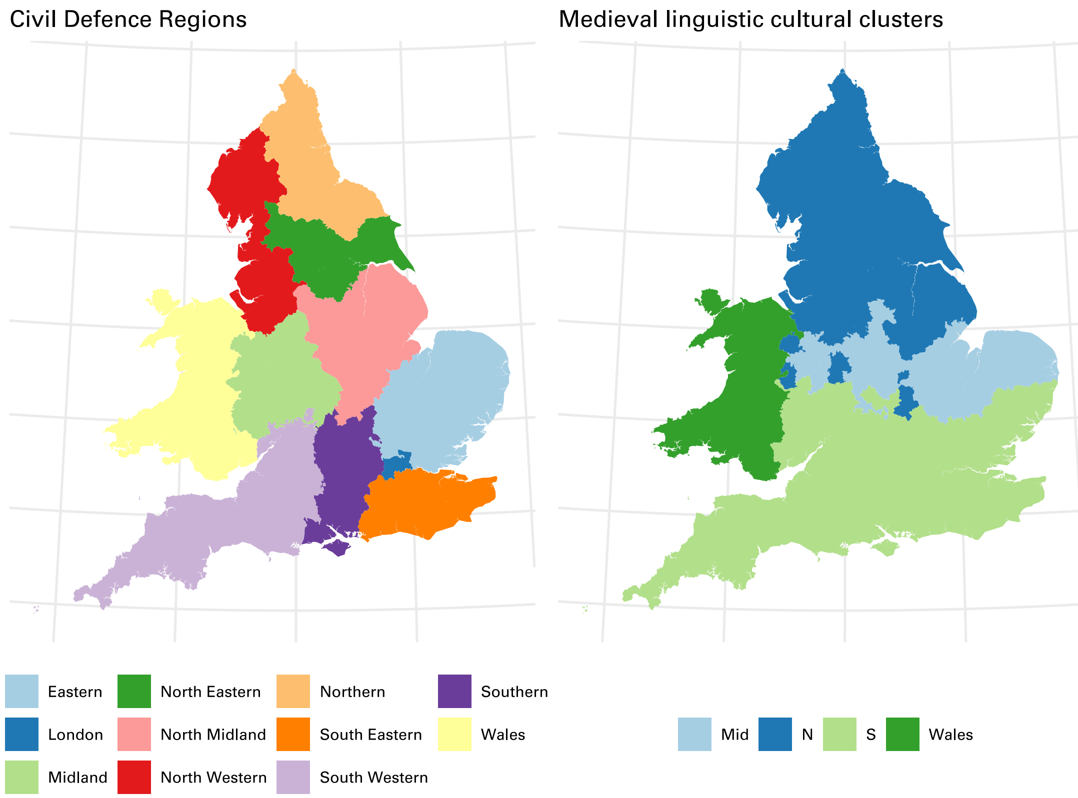

Figure H.5: Alternative cultural clusters

Broadly speaking, our results with these alternative clusters are similar to those using clusters based on historical surnames. Tables H.4 and H.5 show the reduced form analyses with these new clusters, and Tables H.6 and H.7 show the relationship between changes in predicted and actual cultural choices. The results for the civil defence regions are very close to our baseline estimates. In a few cases the results do not hold up with the linguistic clusters, but this is likely attributable to there being fewer clusters, which means there is less variation. For instance, the TSLS effect of migration to a cluster on name scores is smaller and very imprecisely estimated (Table H.5). That is to be expected as the instrument identifies off the migration pull to different clusters due to coal deposits, which affects fewer people if the clusters are larger.

In the structural analyses, the different-sized clusters lead to differences in quantitative magnitudes, and to the effects on the home culture, but other effects are qualitatively similar. Tables H.8 and H.9 show the same counterfactual outputs as Table 6, using the alternative clusters. The effects on the popularity of the Southeastern, Southern, and London cultures, are very similar across these specifications, which increases our confidence in our conclusion that

	ln share migrating			ln share names	
	(1)	(2)	(3)	(4)	(5)
name score	2.409	0.741			
	(0.092)	(0.030)			
recentered coal-predicted					
ln share migrating			4.421		
			(0.608)		
ln share migrating				0.139	0.157
				(0.004)	(0.011)
Name x district FE	x	x			
District FE			x	x	x
Cluster FE	x		x	x	x
Cluster x district FE		x			
Model	OLS	OLS	First Stage	OLS	TSLS
First stage F-stat					52.9
N	1033825	1033825	8990	8990	8990
R^2	0.240	0.919	0.262	0.941	0.940

This table presents evidence of the relationship between cultural naming choices and migration, and replicates Table 2 using civil defence regions in the place of clusters based on surnames before 1800. Models (1) and (2) are estimated at the name-district-cultural cluster level: the dependent variable is the log share of people with a given name born in a given district migrating to a district in a given cultural cluster. The independent variable is the name score for that name for the destination cultural cluster. Both models include fixed effects for the name-district of birth combination, (1) includes fixed effects for the destination cluster, (2) interacts these with the district of birth. (1) and (2) are weighted by the number of people with each name born in each district. Models (3)–(5) are estimated at the district-cultural cluster level. In (4) and (5) the dependent variable is the log share given names most associated with the cultural cluster, and the independent variable is the log share of individuals migrating from the district to that cluster. In (5) this is instrumented for with the log share of migrants predicted by the location of coal deposits in a gravity model, recentered following Borusyak and Hull (2023). We permute the vector of coal deposits across district, calculate predicted log share of migrants under each permutation, and substract the mean of this from the instrument. (3) shows the first stage. (3)–(5) all include district and cluster fixed effects, and are weighted by the number of individuals with name scores born in each district. Standard errors clustered by district in parentheses.

Table H.4: Relationship between migration and naming patterns, with civil defence regions

	ln share migrating			ln share names	
	(1)	(2)	(3)	(4)	(5)
name score	2.398	0.587			
	(0.119)	(0.030)			
recentered coal-predicted					
ln share migrating			1.925		
			(0.544)		
ln share migrating				0.127	0.039
				(0.004)	(0.032)
Name x district FE	x	x			
District FE			x	x	x
Cluster FE	x		x	x	x
Cluster x district FE		x			
Model	OLS	OLS	First Stage	OLS	TSLS
First stage F-stat					12.5
N	565872	565872	3292	3292	3292
R^2	0.488	0.965	0.509	0.889	0.845

This table presents evidence of the relationship between cultural naming choices and migration, and replicates Table 2 using linguistic clusters in the place of clusters based on surnames before 1800. Models (1) and (2) are estimated at the name-district-cultural cluster level: the dependent variable is the log share of people with a given name born in a given district migrating to a district in a given cultural cluster. The independent variable is the name score for that name for the destination cultural cluster. Both models include fixed effects for the name-district of birth combination, (1) includes fixed effects for the destination cluster, (2) interacts these with the district of birth. (1) and (2) are weighted by the number of people with each name born in each district. Models (3)–(5) are estimated at the district-cultural cluster level. In (4) and (5) the dependent variable is the log share given names most associated with the cultural cluster, and the independent variable is the log share of individuals migrating from the district to that cluster. In (5) this is instrumented for with the log share of migrants predicted by the location of coal deposits in a gravity model, recentered following Borusyak and Hull (2023). We permute the vector of coal deposits across district, calculate predicted log share of migrants under each permutation, and substract the mean of this from the instrument. (3) shows the first stage. (3)–(5) all include district and cluster fixed effects, and are weighted by the number of individuals with name scores born in each district. Standard errors clustered by district in parentheses.

Table H.5: Relationship between migration and naming patterns, with linguistic clusters

	Observed cultural change ($\Delta \ln \sigma$)			
	(1)	(2)	(3)	(4)
Predicted cultural change ($\Delta \ln \sigma$)	0.290	0.278	0.166	0.065
	(0.011)	(0.012)	(0.014)	(0.013)
District FE		x		x
Cluster FE			x	x
N	8999	8999	8999	8999
R^2	0.077	0.111	0.645	0.686

This table shows OLS estimates at the district-by-cluster level, replicating Table 5 using civil defence regions. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination real wages v_d^θ, starting populations, and migration costs and the observed value for those born 1861–1895. The dependent variable is the change between the observed value for those born 1841–1860 and those born 1861–1895. Model (2) adds district fixed effects, (3) adds cluster fixed effects, (4) adds both. Standard errors clustered by district in parentheses.

Table H.6: Relationship between the change in log cultural choice shares, σ, relative to the 1851 counterfactual and relative to the 1841–1860 cohort, with civil defence regions

	Observed cultural change ($\Delta \ln \sigma$)			
	(1)	(2)	(3)	(4)
Predicted cultural change ($\Delta \ln \sigma$)	0.549	0.609	0.085	−0.070
	(0.025)	(0.027)	(0.013)	(0.012)
District FE		x		x
Cluster FE			x	x
N	3306	3306	3306	3306
R^2	0.199	0.247	0.802	0.867

This table shows OLS estimates at the district-by-cluster level, replicating Table 5 using linguistic clusters. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination real wages v_d^θ, starting populations, and migration costs and the observed value for those born 1861–1895. The dependent variable is the change between the observed value for those born 1841–1860 and those born 1861–1895. Model (2) adds district fixed effects, (3) adds cluster fixed effects, (4) adds both. Standard errors clustered by district in parentheses.

Table H.7: Relationship between the change in log cultural choice shares, σ, relative to the 1851 counterfactual and relative to the 1841–1860 cohort, with linguistic clusters

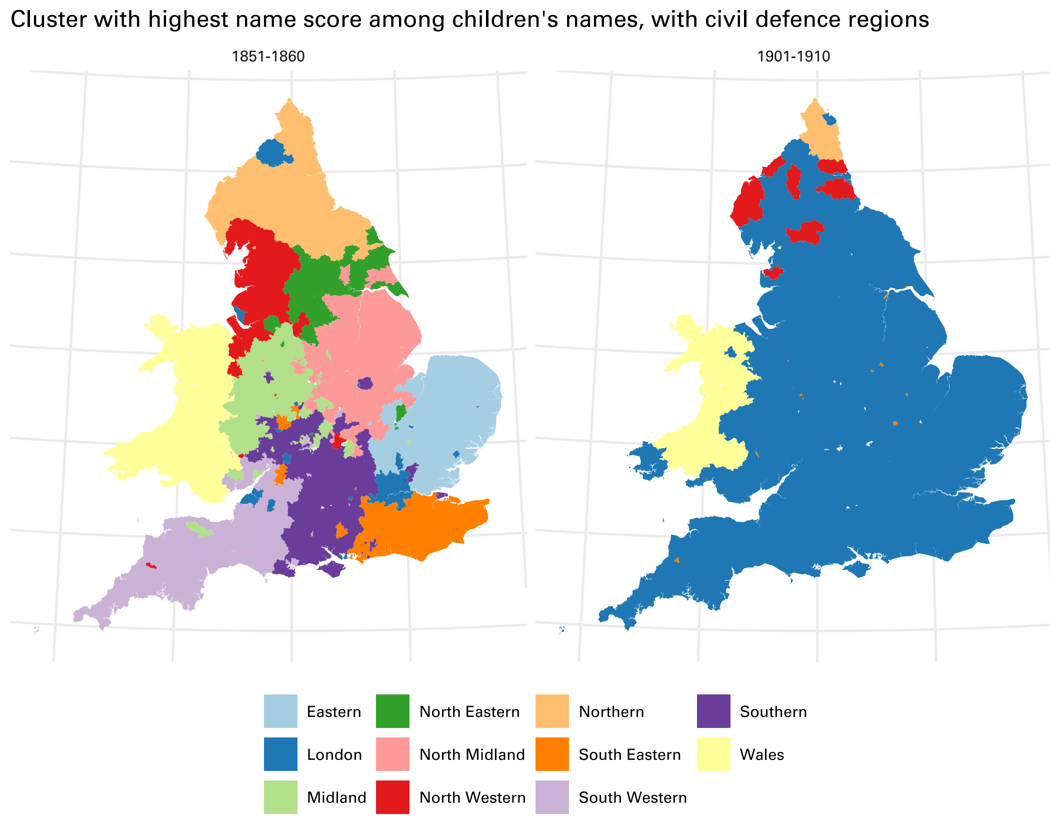

Figure H.6: Changing popularity of cultures, using civil defence regions as clusters

economic changes during the 1851–1911 period bolstered the hegemonic culture. The spatial patterns of counterfactual changes in adoption of these cultures are also similar (Figures H.8 and H.9). Our estimates for the effects on the popularity of home cultures differ in their aggregate magnitudes but follow similar spatial patterns (Figures H.10 and Figures H.11).

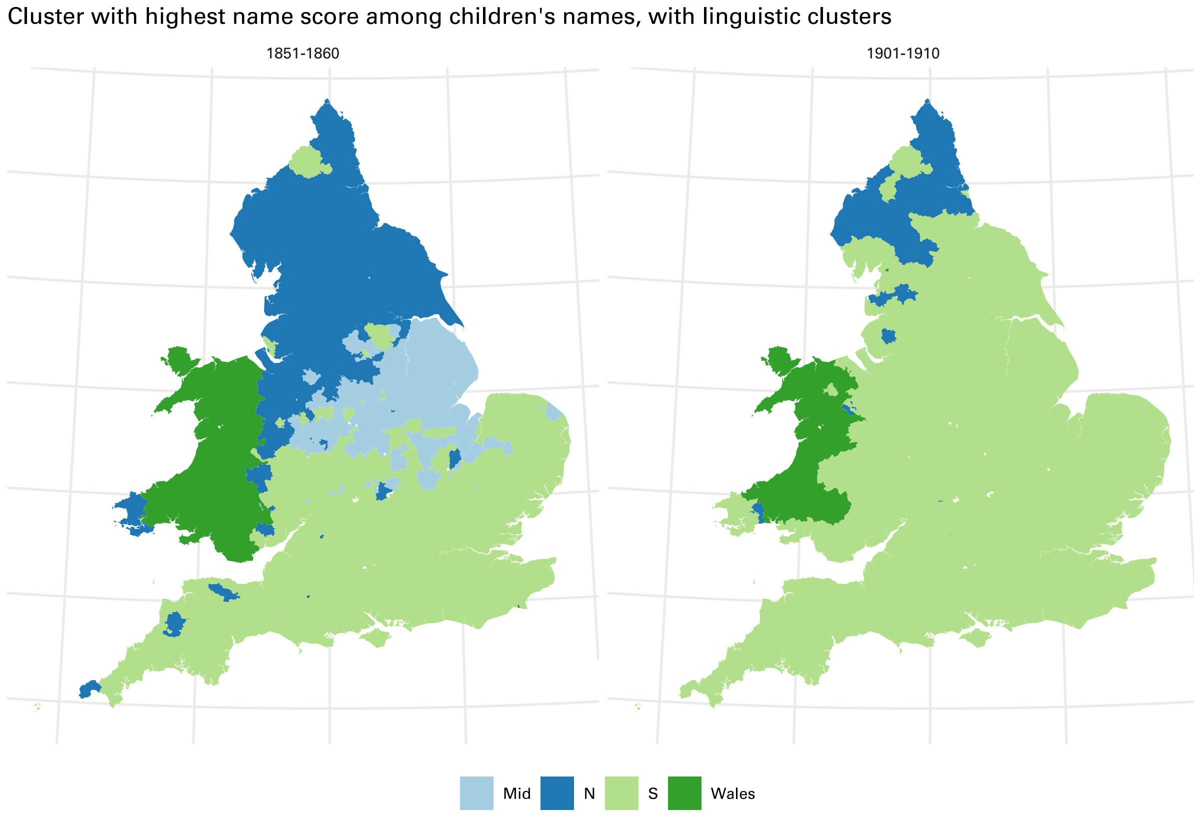

Figure H.7: Changing popularity of cultures, using linguistic clusters

Table H.8: Counterfactual estimates with civil defence regions

	Home culture pop.		London culture pop.		Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	fixing pop.
All 1851 parameters	5.926	7.309	-43.122	-40.602	-35.698	-31.062
1851 real wages, v_d^θ	-5.401	-5.401	-42.189	-42.189	-1.561	-1.561
1851 starting populations	-13.544	-5.174	-29.714	-22.306	-0.727	0.580
1851 migration costs	22.241	22.241	-3.300	-3.300	-31.383	-31.383

This table shows the same counterfactual outputs as Table 6, using civil defence regions in place of clusters based on surnames before 1800.

Table H.9: Counterfactual estimates with linguistic clusters

	Home culture pop.		S culture pop.		Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	fixing pop.
All 1851 parameters	3.668	-2.359	-20.859	-18.306	-35.218	-30.776
1851 real wages, v_d^θ	-9.744	-9.744	-26.941	-26.941	-1.597	-1.597
1851 starting populations	-1.544	-4.509	-14.534	-7.274	-0.690	0.605
1851 migration costs	10.277	10.277	-0.328	-0.328	-31.099	-31.099

This table shows the same counterfactual outputs as Table 6, using linguistic clusters in place of clusters based on surnames before 1800.

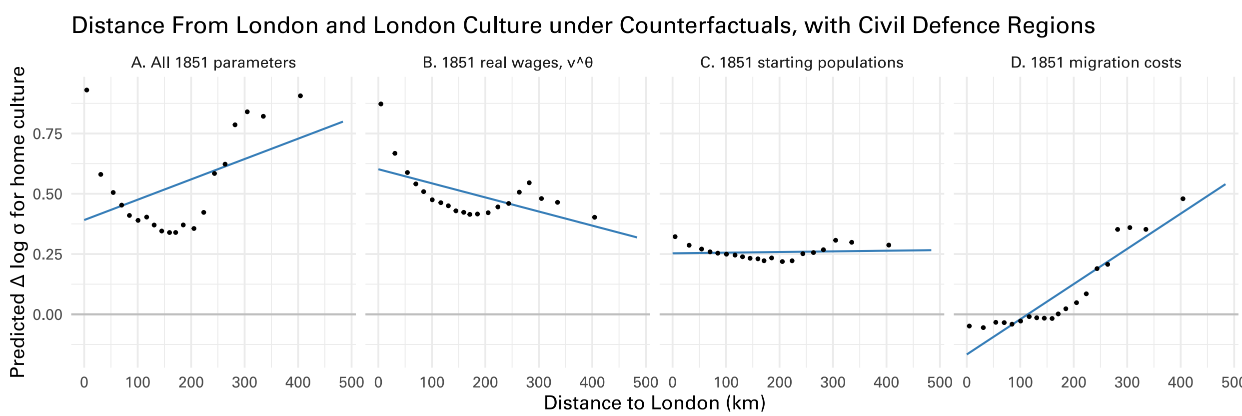

Figure H.8: Distance to London, coal, and the changing popularity of the London culture over the late 19th century, with Civil Defence Regions

This figure replicates Figure 10 using Civil Defence Regions as cultural clusters, and the London region in place of the Southeast region. The figure shows the predicted change in the log share allocated names most associated with the London culture, across different counterfactual scenarios. The y axis is the observed log share minus the counterfactual log share, the x axis distance from the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B only 1851 real wages, C 1851 starting populations, and D 1851 migration costs.

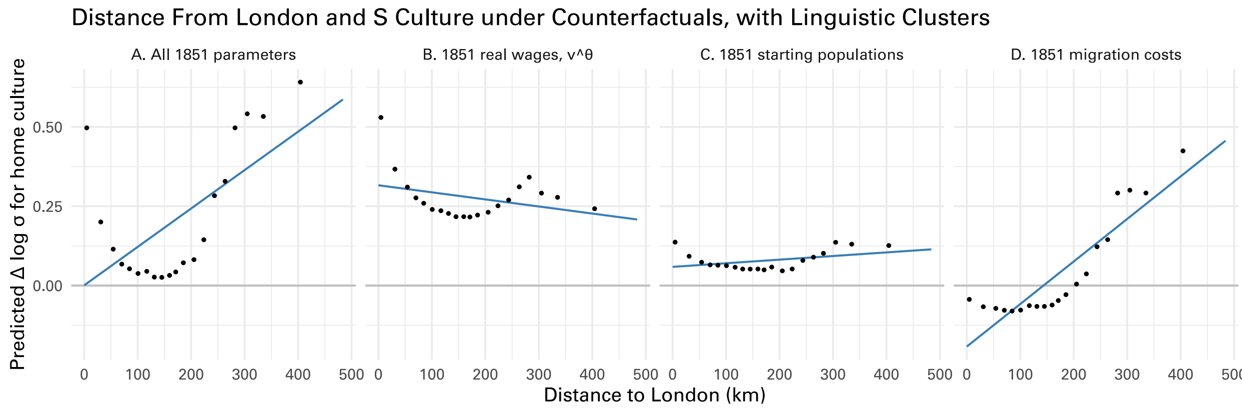

Figure H.9: Distance to London, coal, and the changing popularity of the South English culture over the late 19th century, with linguistic clusters

This figure replicates Figure 10 using linguistic clusters, and the South English cluster, containing London, in place of the Southeast English cluster. The figure shows the predicted change in the log share allocated names most associated with the Southern culture, across different counterfactual scenarios. The y axis is the observed log share minus the counterfactual log share, the x axis distance to the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B only 1851 real wages, C 1851 starting populations, and D 1851 migration costs.

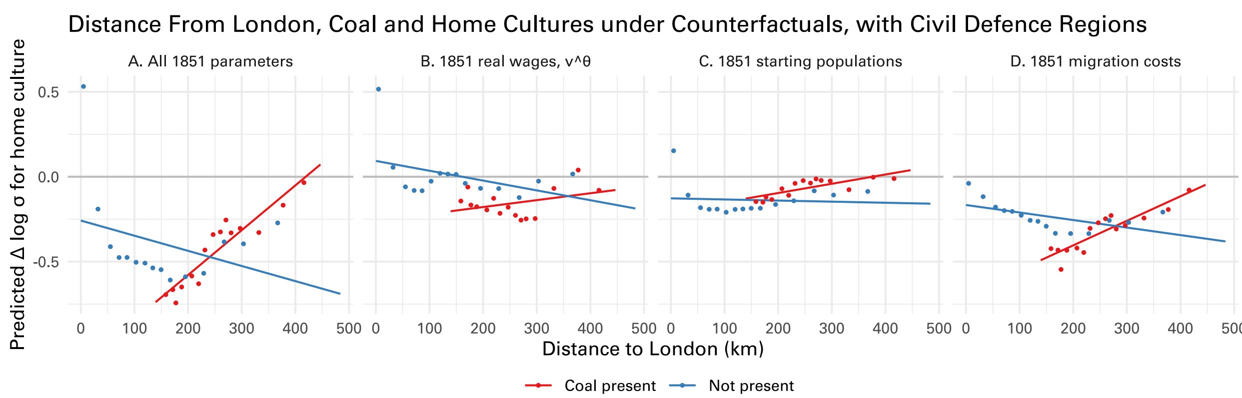

Figure H.10: Average distance, coal, and the changing popularity of home cultures over the late 19th century, with Civil Defence Regions

This figure replicates Figure 11 using Civil Defence Regions as cultural clusters. The figure shows the predicted change in the log share allocated names most associated with the home culture, across different counterfactual scenarios, subset by whether the district contains a coal deposit. The y axis is the observed log share minus the counterfactual log share, the x axis distance to the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B only 1851 real wages, C 1851 starting populations, and D 1851 migration costs.

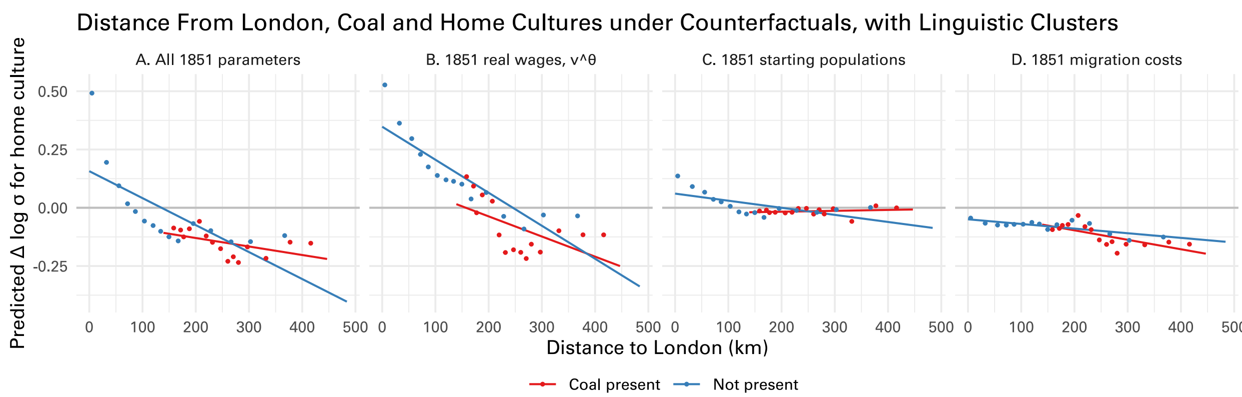

Figure H.11: Average distance, coal, and the changing popularity of home cultures over the late 19th century, with linguistic clusters

This figure replicates Figure 11 using linguistic clusters. The figure shows the predicted change in the log share allocated names most associated with the home culture, across different counterfactual scenarios, subset by whether the district contains a coal deposit. The y axis is the observed log share minus the counterfactual log share, the x axis distance to the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B only 1851 real wages, C 1851 starting populations, and D 1851 migration costs.