INDUSTRY AND IDENTITY
The Migration Linkage Between Economic and Cultural Change in 19th Century Britain

Vasiliki Fouka*
Theo Serlin†

August 2023

Abstract

We study the role of migration in changing the cultural map of England and Wales during the Second Industrial Revolution. In the latter half of the 19th century, Britain underwent a shift in the spatial pattern of economic activity and significant cultural convergence towards the culture of London. Using rich microdata on individuals' names and migration decisions, we document that this homogenization varied across regions and that heterogeneity was mediated by migration patterns. To characterize the heterogeneity, we develop and estimate a quantitative spatial model in which individuals choose migration destinations and cultural identities driven by both economic and cultural considerations. The model indicates that industrialization dilutes local cultures in central hubs which attract migrants but preserves it in peripheral areas by reducing the incentives to out-migrate. Our results provide an explanation for the persistence of local identities in peripheral regions that develop economically and revise the prevailing notion in the modernization literature, which mainly emphasizes the homogenizing effect of labor migration.

*Stanford University, vfouka@stanford.edu
†Stanford University, tserlin@stanford.edu

1 INTRODUCTION

Industrialization had far-reaching consequences, transforming societies from rural and agricultural to urban and industrial. Among these implications were changes in group identities. Classical scholarship from Karl Marx to Ernest Gellner links the development of industry to the reconfiguration of individuals' cultural ties and the emergence of new group loyalties. On the one hand, industrialization is thought to have promoted the breakup of local and particularistic cultural affiliations (Marx and Engels, 1978) and their consolidation into overarching identities that led to the formation of modern nations (Gellner, 2006). On the other, uneven industrial development is also believed to have triggered forces leading to differentiation and the strengthening of pre-existing cultural differences (Gellner 1964, 166–168; Deutsch 1969, 20–25; Hechter 1977).

Most of this literature emphasizes migration as one of the central mechanisms linking 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 favors the adoption of a single identity to serve as a coordinating device. That identity may ensue from migrants adopting the culture of the destination (Gellner, 2006, 59), or—as in the case of peripheral industrial hubs in 19th century France—that of a politically dominant center which acts as a focal point among people of diverse origins (Weber, 1976, 78). Economic migration can thus both reinforce local cultures, or promote their disappearance. Seminal theories of industrialization and identity change recognize a role for migration, but do not draw out its full implications for the cultural landscape that follows industrial change. When does migration driven by changes in economic fundamentals favor homogenization and when does it promote cultural retention?

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. We use rich census microdata on households' migration and cultural choices to study how the changing geography of industry 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 and its connection to changing economic activity. 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 observe two patterns. First, between 1851 and 1911, England's cultural landscape underwent significant homogenization. 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, there was considerable spatial heterogeneity in cultural homogenization. Regions geographically further from London experienced greater assimilation into the Southeastern culture. Within more peripheral regions, industrial hubs assimilated at a slower rate, but in the interior the opposite was true.

We next provide descriptive evidence that internal mobility was a crucial link between industrialization and these heterogeneous patterns of cultural change. During the period under study migration increased, with out-migration rates from some regions rising up to 15 percentage points. Migration patterns were influenced by industrial development. People left agricultural areas and moved towards regions with high manufacturing growth. These migration choices were both affected by culture and in turn influenced cultural choices. Regions that experienced increases in the rate of out- or in-migration assimilated into the Southeast cluster at a faster rate. People were more likely to migrate to culturally similar areas, as indicated by their first names. 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.

Together, these stylized facts suggest that the relationship between industrialization, migration and cultural change varies across contexts and complicate the conventional wisdom in the modernization literature, according to which labor migration was a purely homogenizing force. Industrializing areas did indeed experience higher in-migration, a force pushing towards cultural change and coordination on an overarching identity. At the same time, these locations experienced lower out-migration, a force in the direction of cultural retention. The resulting outcome should depend on where industrialization was located. Geographic location relative to other industrializing centers affects where in-migrants will be drawn from and whether they will mainly originate from different cultures promoting homogenization or from the local culture promoting retention. Furthermore, because people have a preference for being around others of the same culture, whether in-migrants assimilate into the culture of the destination, or whether those in the destination adopt the culture of the in-migrants will depend on coordination dynamics.

To systematically describe heterogeneity in the effects of industrialization on culture and explore the implications of coordination in migration and cultural choice, we build a theoretical model, which incorporates the stylized facts we document. In the model, 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. 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 geographical 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. The model also predicts changes in cultural choices due to economic developments over the 1851–1911 period correlated with observed changes.

We use the model to conduct counterfactual exercises for migration and cultural choices under different values of economic fundamentals. The model rationalizes the empirical patterns we observe. If real wages had remained constant in their 1851 level, the prevalence of the culture of Southeast England in 1911 would be 5% lower. Though the effect on net is homogenization, there is substantial variation. The model allows us to quantify this heterogeneity, by examining how cultural choice in each location changes in response to changes in the real wage in that location, holding other parameters constant. 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. Finally, we also ask how economic outcomes would change if individuals had different cultural preferences over locations. Counterfactual simulations under different place-specific preferences suggest that culture-specific attachments drive individuals away from locations that could afford them higher real wages. Complete assimilation into the culture of the Southeast, by reallocating migrants across space, increases average income by 0.5–1.6%, depending on different estimates of the wage-migration elasticity.

This paper makes four main contributions. First, we demonstrate theoretically and empirically how internal mobility serves as a link between economic and cultural change. In so doing, we build on and empirically validate insights in the modernization literature that connect industrialization to the formation of overarching identities and the process of cultural convergence. While these theories identify multiple channels mediating this connection, such as education (Gellner, 1964), cultural diffusion through print media (Anderson, 1983), or the constructions of infrastructure networks across markets (Weber, 1976), many underscore labor migration's unique role.[1] We explicitly model this role in a setting in which both migration and culture are individual choices, made interdependently as people attempt to coordinate with others. Under these assumptions, we find that migration leads to both homogenization and retention of cultural difference, depending not only on locations' economic fundamentals and local culture, but also those of other locations around them.

Second, 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.

Third, 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

[1] More recent work like Green (2019) and Green (2022) also links industrialization to assimilation into larger cultures. Similar to our paper, these studies emphasize economic incentives and bottom-up processes of assimilation rather than state-led nation-building.

second half of the 19th century. Recent work in urban and international economics has 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.

Fourth, 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. In Section 2, we describe our data and the methods we use to quantify early local cultures and measure cultural change in England and Wales. In Section 3, we present descriptive patterns on industrialization, cultural homogenization and migration during the second Industrial Revolution. Section 4 introduces our theoretical model, estimation strategy and results and validation of recovered structural parameters. In Section 5, we conduct counterfactual exercises to quantify the role of changes in economic fundamentals for cultural change. Section 6 concludes.

2 DATA AND MEASUREMENT

This section introduces our data, and the methodologies we use to measure the distribution of cultures before the Second Industrial Revolution and subsequent cultural change.

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, the British census recorded granular information on birthplaces, which allow us to trace out lifetime migration. Second, 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.

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.[2] The advantage of focusing on registration districts is that other social and economic data from this period, 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).

2.1 *Cultural Clusters Before the Second Industrial Revolution*

Our analysis focuses on how the economic changes brought about by the Second Industrial Revolution influenced the popularity of different cultures. As a prerequisite for analyzing cultural change, we need an allocation of districts to cultural clusters prior to the developments we analyze.

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

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

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.

This combination of surname data before the Second Industrial Revolution and a spectral clustering algorithm generated clusters that correspond to historical regions. For instance the W region of Figure 1 corresponds with 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 A-1). Comparing pairs of individuals resident in the same parish, those born in different clusters were less likely to be married (Table A-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. However we do not believe these clusters represent the only way to divide up England and Wales according to culture, and in Appendix D we examine alternative cultural clusters and verify that our results are robust to changing cluster boundaries.

2.2 *Measuring Cultural Choices*

Much of our analysis focuses on cultural choices, which we proxy using names. In the main analysis we use data on the naming and migration histories of those born between 1861 and 1895, who would have been adults in 1911. To allocate names to cultural groups, we use data on the previous generation, those born between 1841 and 1860. Using the frequencies of names among those born 1841–1860 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 cultural cluster.[3] 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).

[3]We standardize name spellings using Metaphone codes.

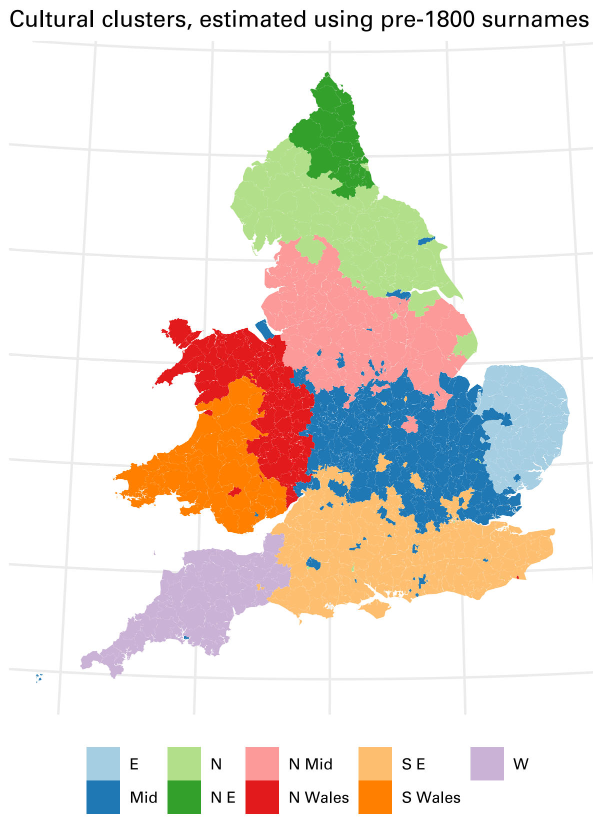

Figure 1: Map of clusters

Clusters from running spectral clustering algorithm on surname frequencies for those born before 1800 present in the 1851 census

Given name scores calculated from those born between 1841 and 1860 allocated to these cultural clusters, we calculate origin-by-destination-by-culture migration flows using those born between 1861 and 1895, which we use for out main analysis. This age bracket corresponds to adults of working age in the 1911 census. In supplementary analyses at the district level, we use the popularity of culture-specific names among those born 1851–1860 and 1901–1910. We also use the census to measure intermarriage between those born in different clusters, and the growth of manufacturing across regions over time.

3 Economic and cultural change in the late 19th century

This section discusses economic and cultural change in England and Wales over the second half of the 19th century. The spatial patterns of both economic activity and cultural affiliation changed dramatically over this period. The development of new industries stimulated economic growth outside the heartland of the First Industrial Revolution. Cultural choices became more homogeneous, though homogenization proceeded unevenly across space. Migration provides a plausible mechanism connecting these two developments. People migrated at higher rates, especially to the regions that were rapidly industrializing, and made cultural choices in anticipation of future migration.

In 1851, manufacturing in Britain 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.[4]

The growth of new industries altered the spatial distribution of economic activity. Figure 2 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 experienced rapid growth, in part due to new industries locating close to investors in the City of London (see also Geary and Stark 2015).

As the Southeast grew economically, its culture grew in popularity. Figure 3 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 most notable development is cultural homogenization. The culture of the Southeast of England increased in popularity relative to other cultures; entire cultures disappear 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. The rapid increase in the Southeast culture is suggestive of a coordination dynamic of the type studied by Laitin (1994). If there are benefits of adopting the same culture as one's neighbor, we might expect people across regions to coordinate on one culture. It is also

[4]These figures were calculated from I-CeM census data (Schürer and Higgs, 2014).

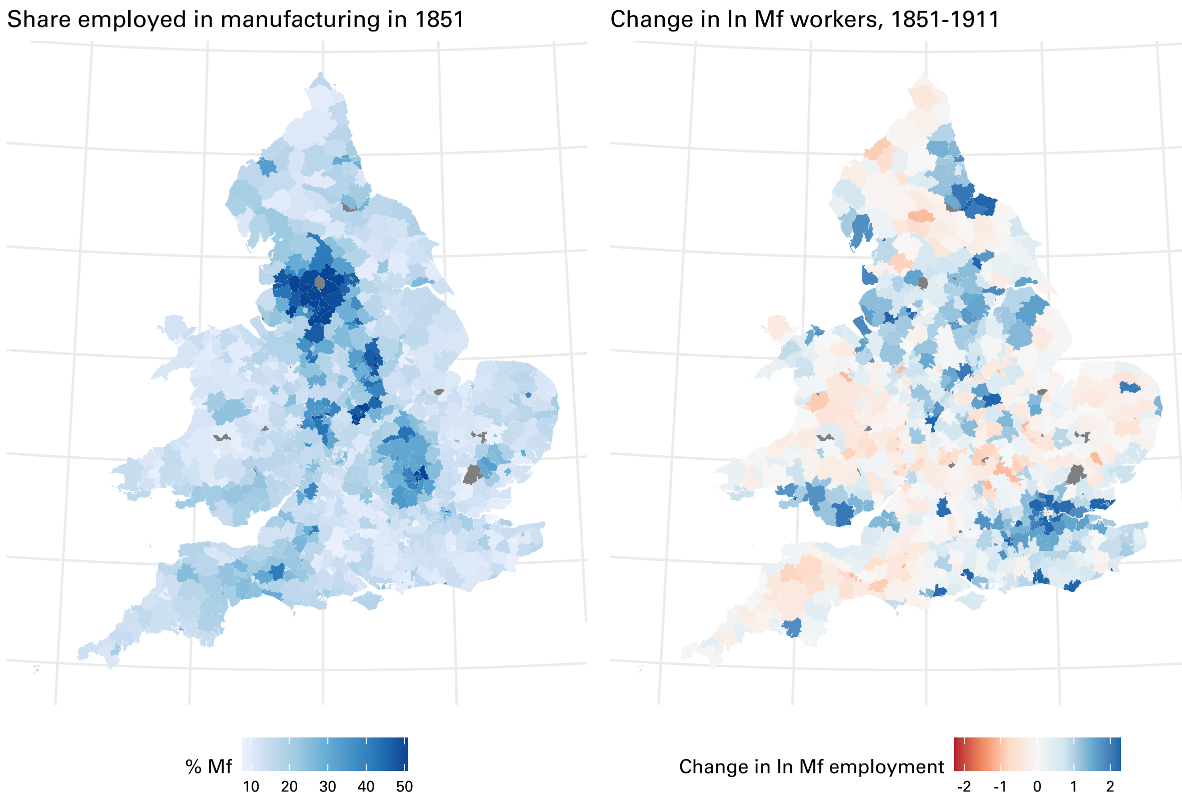

Figure 2: Industrialization in 1851, and growth over the 1851–1911 period

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.

similar to the adoption of Paris French in industrializing regions of France that saw diverse in-migration from a other regions (Weber, 1976).

Figure 4 shows average name scores for home cultures and for the Southeast culture, among children born in different clusters present in the 1911 census, and presents a similar pattern to Figure 3. The average name score for all cultures except the Southeast declined, and in each cluster average name scores for the Southeast increased. The regions that do not appear to have been absorbed into the Southeast in Figure 3 did also experience large increases in name scores for the Southeast, but from a lower base. While we are wary of interpreting these changes in naming patterns as being entirely due to cultural homogenization, we do observe a similar pattern in rates of inter-cluster marriage. The left panel of Figure 5 shows the share of married people born in a given cluster in a given year of the census married to people born in the same cluster. Rates of intermarriage increased for all clusters, but much less for the Southeast than for other cultures.

Homogenization into the Southeast culture was heterogeneous. In both relative and

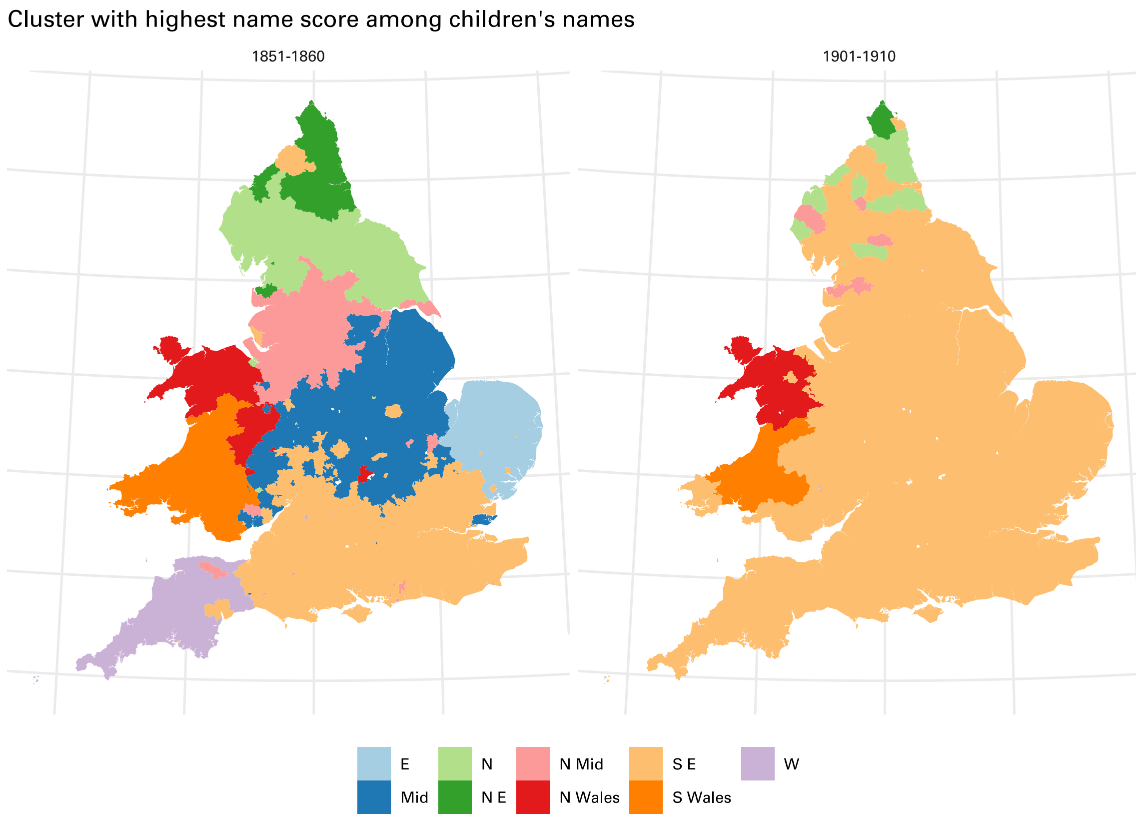

Figure 3: Changing spatial distribution of culture-specific names

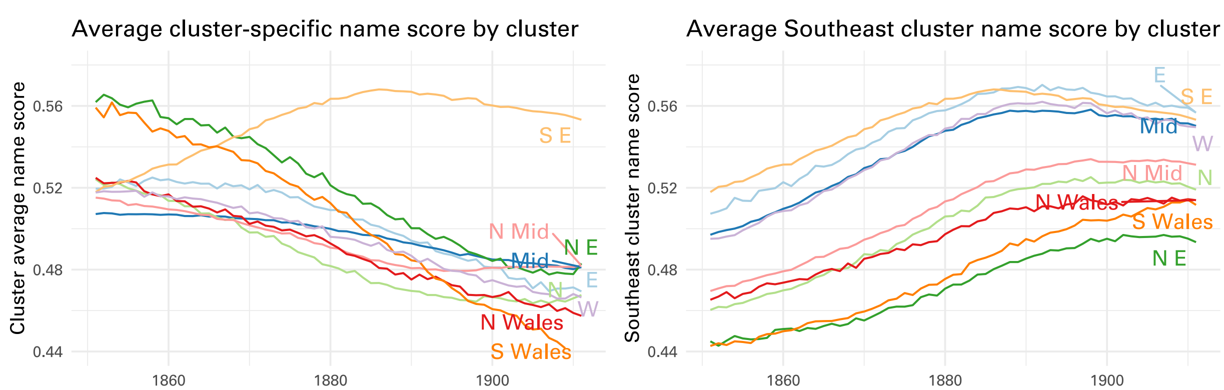

Figure 4: Decrease in average name scores for clusters other than Southeast England, increase in name scores for Southeast England in all regions

Left panel shows the average name score for the home cluster among those present in the 1911 census born in a given cluster in a given year; right panel shows average name scores for the Southeast England cluster

absolute terms, peripheral districts, for instance those in the Northeast and Wales, experienced larger declines in name scores for their home cultures, and increases in average name scores for

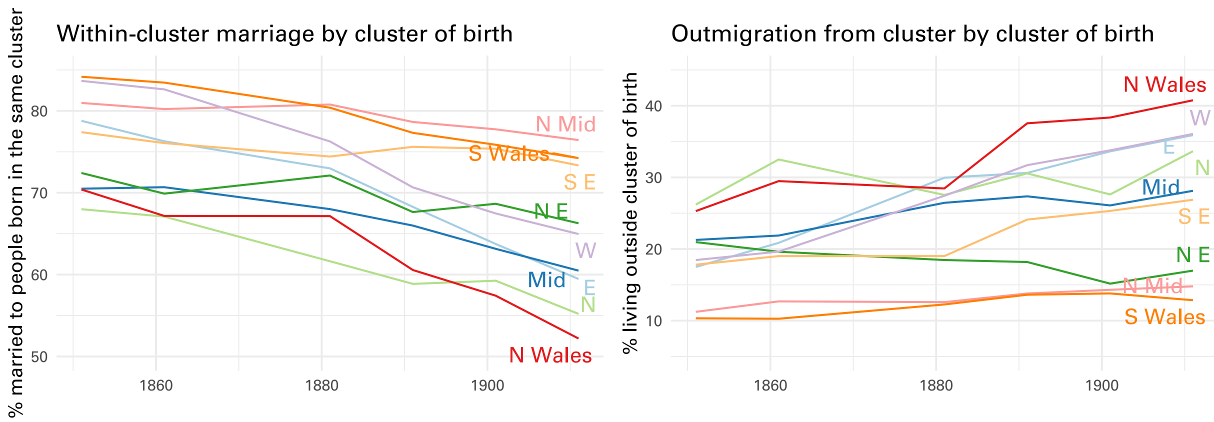

Figure 5: Increasing marriage outside the cluster and migration

Left panel shows the share of married individuals born in each cluster married to someone born in the same cluster; right panel shows the share born in a given cluster resident outside the cluster

the Southeast culture (Figure 4). Figure 6 shows the spatial distribution of the change in the log share allocated names most associated with the Southeast culture, between the 1851–1860 and 1901–1910 cohorts. Locations further from London experienced larger relative increases in the popularity of the Southeast culture. In general, rural areas, especially in Wales, experienced the largest relative increases. Within regions, the relationship between industrialization and assimilation into the Southeast varied. The South Wales coalfield assimilated at a slightly slower rate than more rural parts of Wales (Figure A-1 provides a map of districts containing coal deposits). In contrast, industrial parts of Lancashire experienced some of the strongest assimilation into the Southeast.

3.1 *Migration Linkages Between Industrialization and Cultural Change*

These spatial patterns are consistent with changing patterns of migration driving cultural homogenization. If people migrate in part for economic reasons, and in part to be with people of the same culture, and choose culture anticipating future migration, then industrialization, by changing where people migrate, can influence cultural choices. The growing appeal of the Southeast for migrants should thus increase the popularity of the Southeast culture. Declining costs of migration should have the largest effects in places on the periphery that previously had low rates of migration to the Southeast and to other interior regions that were culturally similar to the Southeast. Local industrialization provides an alternative to out-migration, which may reduce assimilation into the Southeast. However, it may also attract in-migrants, who may bring other cultures with them. On the periphery, it is likely that reductions to out-migration dominate increases in in-migration and serve to preserve local culture, while in locations closer to the interior, in-migration responses to local growth are stronger than

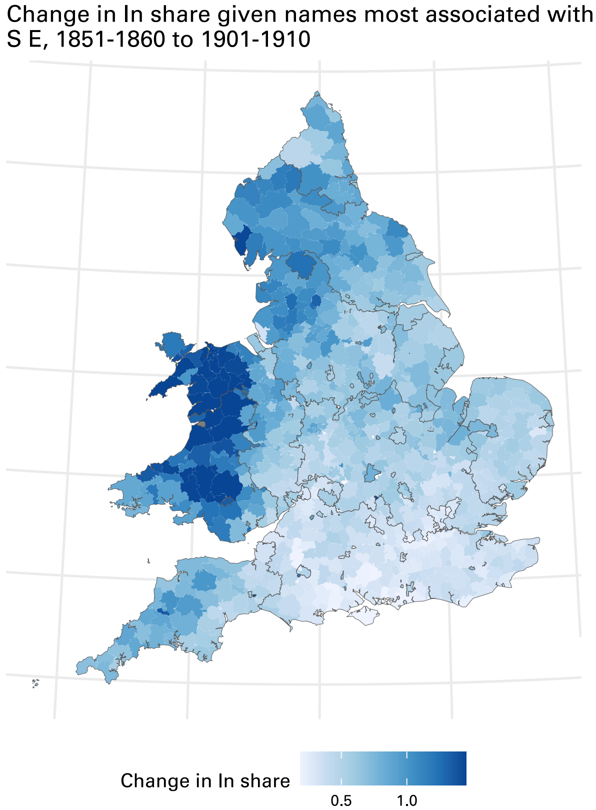

Figure 6: Uneven assimilation into the Southeast

out-migration responses.

Migration increased over this period, making it more plausible that migration linked economic and cultural change. The right panel of Figure 5 shows rates of out-migration from different clusters. The overall level of migration increased, and the regions from which people migrated changed. The increase in the level of migration implies that the cost of migrating decreased. Even growing regions like South Wales experienced higher rates of out-migration. People were less likely to migrate from, and more likely to migrate to, industrializing regions. The most conspicuous increases in migration were from regions like North Wales, the West Country (W) and East Anglia (E) that did not meaningfully industrialize. While decreasing costs increased the level of migration, the changing pattern of industrialization created incentives for people to migrate to a different set of regions. In Appendix B, we use a gravity model to separate the contributions to increased migration of changes to the pull of particular destinations and to migration costs. While both contributed to the increase, the contribution of reduced migration costs was larger.

Migration and cultural change were not just coterminous but inter-related. Table 1 shows

	Δ ln share S E culture			Δ out-mig.	Δ in-mig.
	(1)	(2)	(3)	(4)	(5)
Δ ln out-migrants / pop	0.128				
	(0.050)				
Δ ln in-migrants / pop		0.066			
		(0.018)			
Δ ln Mf workers			0.062	−0.132	0.331
			(0.024)	(0.034)	(0.085)
N	810	809	809	816	815
R^2	0.029	0.062	0.048	0.060	0.065

This table presents the results of district-level regressions linking cultural change, migration and industrialization. In models (1)–(3), 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 (4) the change in the log number of people over 16 born in the district living outside the cluster, divided by the number born in the district, between 1851 and 1901, in (5) the change in the log number over 16 born outside the cluster living in the district, divided by the number born in the district. The independent variable in (1) is the dependent variable in (4). In (2), the dependent variable in (5). In (3)–(5), the change in the log number of manufacturing workers, between 1851 and 1901. In (1)–(3) observations are weighted by the number allocated namescores in the 1851–1860 cohort, in (4) and (5) by 1851 population. Robust standard errors in parentheses.

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

the results of regressions of the change in the log share given names most associated with the culture of the Southeast of England between the 1851–1860 cohorts and the 1901–1910 cohorts—the variable mapped in Figure 6—against changes in out-migration, in-migration and industrialization over the same period. Increases in both out-migration and in-migration are associated with increases in the popularity of the Southeast culture. Table A-3 shows that in- and out-migration was associated with declines in the home culture.

Industrialization is not robustly correlated with cultural change.[5] This null result makes sense because industrialization both encourages in-migration and discourages out-migration, as shown in models (4) and (5). Whether local industrialization increases or decreases the appeal of a particular culture will depend on geographic factors that condition which of these effects dominate. We return to this question in Section 5.

People were more likely to migrate to locations that were culturally-similar to them. The first two columns of Table 2 show 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,

[5] Though in Table 1 model (3) the association is statistically significant, this relationship is sensitive to specification choices.

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.

It was not just that cultural choices were predictive of migration; migration opportunities influenced cultural choices. Columns (3)–(5) examine the relationship between the log 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. 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 predicted by coal deposits in a gravity model. We first regress origin-by-destination log migration flows against origin and destination fixed effects, log distance and in indicator that the origin and destination are the same. Following best practices in economic geography we estimate the model by Poisson pseudo-maximum likelihood (Silva and Tenreyro, 2006; Fally, 2015). We then predict the destination fixed effects with the presence of coal, and use these predicted fixed effects and the estimated distance elasticity to calculate the predicted log share migrating from each district to districts allocated to each cluster. We also recenter this instrument following Borusyak and Hull (2020). 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 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. These adjustments make the exclusion restriction more plausible, as we just identify off heterogeneity in where coal deposits attracted migrants. In both the OLS and TSLS specifications, 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.

These results suggest that industrialization influences migration, and that cultural and migration choices influence one another. If people make cultural choices anticipating where they will migrate, and who else will migrate there, then how a given change in fundamentals affects cultural choices will depend on how these coordination forces play out in equilibrium across locations. Because all these choices are interdependent, if we want to understand the contribution of a given change on observed outcomes we need a methodology that can account for general equilibrium effects.

Not only will a given change in economic fundamentals in one location have rich and

	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

varied effects across other locations; different locations will likely respond differently to a shock of the same magnitude. Whether local economic growth preserves the local culture by deterring out-migration, or undermines it by encouraging in-migration, will depend on the spatial structure of culture and the economy. We would expect the relationship between industrialization and cultural change to be heterogeneous across regions, which warrants a methodology that allows us to quantify heterogeneity. The remainder of the paper sets out a structural model that accounts for general equilibrium effects and helps us analyze heterogeneity across regions.

4 MODEL

This section sets out the theoretical model, and the strategy we use to estimate the key parameters from our data. In the model, agents, who are distinguished by their location of origin and culture, migrate to locations. 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 assume, 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 assign them to one culture, not knowing fully what these individual-level preferences will be. Parents care about the expected welfare of their child under different cultures, but also have their own preferences for specific cultures.

The model generates closed form expressions for culture-origin-destination migration flows and cultural choices at the origin. If we know the values of two elasticities—the elasticity of the culture-specific migration appeal of a place to the number of people of the same culture there, and the elasticity of cultural choices to the expected benefits of migration under that culture—we can back out the various taste parameters. We use an instrumental variables strategy based on the interaction between geographical locations, historical cultures, and coal deposits, to estimate the elasticities. We find that our estimated taste parameters correspond to other measures of cultural proximity.

We treat the observed cultural choices and migration flows as one equilibrium of the model under the observed economic fundamentals. Once we have estimates of the model's theoretical parameters, we can solve the model for counterfactual equilibria under different fundamentals. Doing so allows us to quantify how a change in fundamentals affects outcomes, both in the aggregate and across regions.

4.1 *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: people who are born, assigned culture, and migrate, and their parents, who choose which culture to assign them. When choosing where to migrate, the utility person i of culture k born in location o receives from migrating to destination d is

$$u_{odi}^{k} = v_d \delta_{od} \left(m_d^k\right)^{\frac{\alpha}{\theta}} \xi_d^k \varepsilon_{di}$$

v_d is the real wage in d, an attribute of the location that is experienced equally by members of all cultures.[6] δ_{od} is a distance-related inverse cost of migrating from o to d, and captures the idea, illustrated by gravity models of migration, that people face larger costs of migrating further. m_d^k 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_d^k is a preference shock unobserved by the econometrician, but known to the agents, for location d for all members of culture k. This 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.[7]

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

$$m_{od}^{k} = \frac{\left(v_d \delta_{od} \xi_d^k\right)^{\theta} \left(m_d^k\right)^{\alpha}}{\sum_{j=1}^{N} \left(v_j \delta_{oj} \xi_j^k\right)^{\theta} \left(m_j^k\right)^{\alpha}} n_o^k \tag{1}$$

Where n_o^k is the number of people originating in o with culture k. It also follows from the

[6] 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.

[7] 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.

properties of the Fréchet distribution that expected utility—prior 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:

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

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 choices 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

$$\bar{u}_o^j=\left(\Omega_o^k\right)^{\frac{1}{\theta}}\left(\psi_o^k\right)^{\frac{1}{\varphi\theta}}\iota_i^k$$

where ψ_{ko} is a place-specific bias for a given culture, and ι_i^k is another 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

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

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. As φ tends towards infinity, we would expect all parents in a given location to choose the same culture for their children.

4.2 *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 assumption implies rational expectations on the part of parents: they correctly anticipate the cultural choices of all other parents and the migration opportunities of their children.

Results from Allen, Arkolakis and Li (2020) imply that the equilibrium exists, is unique, and can be computed through an iterative algorithm if $\frac{\alpha}{1-\alpha}\max(2\varphi-1,1)<1$. We present these results in Appendix C. The intuition for this uniqueness result is that if individual culture and migration choices are very responsive to the actions of others—that is, if α and φ are both large—the model becomes a pure coordination game as in the tipping games modeled by Laitin (1994). For lower values of α and φ, heterogeneity in the idiosyncratic preferences of parents for cultures, represented by the Fréchet shocks, acts as a brake on these coordination forces.

4.3 *Estimating Equations*

Using data on culture-specific migration flows, we can estimate this pair of elasticities, and recover the structural variables, in stages.

Taking logarithms of (1) gives

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

Parameterizing migration costs as a log-linear function of geographic distance and an indicator that the origin equals the destination, $\delta_{od}^{\theta}=\text{distance}_{od}^{\beta_1}\exp\left(\beta_2\mathbf{1}_{\{o=d\}}\right)$, we can estimate the following model by Poisson pseudo-maximum likelihood at the culture-origin-destination level

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

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} captures pull factors common to all cultures (v_d^{θ}), and factors specific to each culture ($\left(m_d^{k}\right)^{\alpha}\left(\xi_d^{k}\right)^{\theta}$). We can then regress the destination-by-culture fixed effects against the log of culture-specific migration flows to estimate α and v_d^{θ}:

$$\gamma_d^{k}=\alpha\ln m_d^{k}+\gamma_d+\theta\ln\xi_d^{k} \tag{4}$$

The fixed effect γ_d captures $\theta\ln v_d$. The error term $\theta\ln\xi_d^{k}$ represents unobservable factors that influence culture-specific migration decisions.

Given estimates for δ_{od}^{θ} and the $\left(v_d\xi_d^{k}\right)^{\theta}\left(m_d^{k}\right)^{\alpha}$ bundle, we can calculate Ω_o^{k} according to equation (2).

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

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

This equation implies that regressing the log of the share choosing a given culture in a given location (σ_o^k) against the log of Ω_o^k, calculated for that culture in that location, and a location fixed effect will provide estimates of φ and ψ_o^k.

4.4 *Instrumental Variables Strategies*

4.4.1 α equation

Two endogeneity problems complicate estimating Equation (4). First, culture-specific migration shocks ξ_d^k cause in-migration and so increase m_d^k, so the two should be positively correlated. Second, ξ_d^k influences people's choices of culture, and may also be correlated with ψ_d^k, so the stock of migrants of culture k originating near d should be correlated with ξ_d^k.

We require the following to construct a valid instrument for m_d^k. First, the presence of coal deposits must be correlated with $\theta \ln v_d$ and not with ξ_d^k. The substantive assumption is that the presence of coal influences in-migration through economic incentives which apply equally to all cultures. Second, the culture that we allocate district d to based on surnames before 1800 must be correlated with ψ_d^k. We do not need to require that ψ_d^k and ξ_d^k are uncorrelated, or that a location's surname cluster allocation is uncorrelated with its values of ξ_d^k.

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. This gravity regression gives an alternative set of migration cost coefficients and destination fixed effects. We generate predicted destination fixed effects by regressing the estimated destination fixed effects against an indicator for coal presence. We then use the coal-predicted destination fixed effects, population, and migration cost coefficients to predict migration flows. We assign these migration flows to cultures based on the cultural clusters to which we allocate the origin district based on surnames before 1800. The prediction for m_d^k is therefore the sum of predicted migration flows to d from districts allocated to culture k. This instrument uses variation in the cultural mix across locations due to geography, and does not take into account culture-specific pull and push factors. We then estimate the following equations by two-stage least squares:

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

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

where $\gamma_{k(d)}^k$ is a fixed effect for the surname culture cluster d is located in interacted with the culture in question k. The $\gamma_{k(d)}^k$ fixed effects are important as we might worry that which surname cluster a district is in is correlated with unobservable culture-specific tastes for migrating into that district.

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).

4.4.2 φ equation

The main source of endogeneity in equation (5) is that ψ_o^k increases σ_o^k which increases m_o^k because most people migrate to places close by, which increases Ω_o^k. In words, an exogenous preference for a choosing a given culture in a given location—the error term in equation (5)—causes there to be more people of that culture resident in proximate locations, which influences the migration benefits of belonging to that culture—the independent variable. A secondary source of endogeneity is that place-specific preferences for a given culture ψ_o^k may be correlated with culture-specific preferences for a given place ξ_o^k, which also features in the independent variable Ω_o^k.

As byproducts of estimating α, we have estimates of distance-related migration costs δ_{od}^θ and destination-utilities predicted by coal. We use these to construct an instrument for Ω_o^k:

$$\text{coal-predicted } \Omega_o^k = \sum_{d=1}^{N} \hat{v}_d^\theta \delta_{od}^\theta \mathbf{1}_{\{\text{historic culture}(d)=k\}}$$

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.

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.

We then estimate φ by two-stage least squares as follows:

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

$$\ln \sigma_d^k = \varphi \ln \hat{\Omega}_o^k + \gamma_o + \gamma_{k(o)}^k + \ln \psi_o^k$$

We use our estimate of φ and observed σ_d^k and Ω_o^k to back out ψ_o^k

4.5 *Estimation and Inference*

Our estimation routine proceeds in multiple stages. First, we must allocate individuals to cultures based on name scores. Doing so requires that we calculate name scores, which we do using the names of individuals born 1841–1860. Then we aggregate origin-by-destination-by-culture migration flows using individuals born between 1861 and 1895. 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.

Because the estimation of each parameter uses outputs from other estimations, conventional standard errors will be incorrect. We therefore use a fractional random-weight bootstrap clustered at the district level. On each bootstrap iteration we draw a set of weights for locations, which we use to calculate name scores and estimate the three regressions. We use the distribution of parameter estimates over these iterations to construct confidence intervals.

4.6 *Estimates*

Table 3 shows the results of these estimation routines for α and φ. The two-stage least squares estimates imply $\alpha = 0.26$ and $\varphi = 1.49$, which give an Allen, Arkolakis and Li (2020) bound of $\frac{\alpha(2\varphi-1)}{1-\alpha} = 0.69$, which is less than one and so implies that the equilibrium is unique. Given that the bias from endogeneity on the OLS coefficients should be positive, it is encouraging that the TSLS coefficients are smaller than the OLS coefficients. The F-stat for the instrument for $\ln \Omega_o^k$ is 8.7, which is slightly weak by conventional standards. The Anderson-Rubin test, which is robust to weak instruments, indicates that we can reject the null of $\varphi = 0$ at the 5% level.

4.7 *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 section we

	$\ln m_d^k$	γ_d^k		$\ln \Omega_o^k$	$\ln \sigma_o^k$	
	(1)	(2)	(3)	(4)	(5)	(6)
coal-predicted $\ln m_d^k$	0.107					
	[0.080; 0.126]					
$\ln m_d^k$ $[\alpha]$		0.719	0.260			
		[0.633; 0.766]	[0.077; 0.487]			
Recentered coal-predicted $\ln \Omega_o^k$				0.264		
				[0.096; 0.438]		
$\ln \Omega_o^k$ $[\varphi]$					1.597	1.487
					[1.274; 1.805]	[0.472; 2.650]
Model	First stage	OLS	TSLS	First stage	OLS	TSLS
First stage F-stat			80.406			8.694
AR p-value			0.02			0.016
N	7418	7418	7418	7401	7401	7401
R^2	0.989	0.996	0.994	0.998	0.986	0.986

This table shows OLS and TSLS estimates of α and φ. All models are estimated at the district-by-culture level, weighted by the number of observations in the district. Models (1) and (4) show the first stage, (2) and (5) the OLS and (3) and (6) the TSLS. 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. Bootstrap percentile 95% confidence intervals clustered by district in brackets.

Table 3: Estimates of model elasticities

check whether these estimates correlate with variables not included in the model that should be associated with these economic and taste parameters.

We would expect ξ_d^k and ψ_o^k to be correlated with cultural proximity. Beyond having more people of the same culture, a location may be more attractive to people of culture k if it has amenities specific to that culture, like the same type of church, or if its residents were historically more culturally similar. Similarly, people in a location may be more likely to assimilate into 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 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 choice preferences (ψ_d^k) and migration preferences (ξ_d^k) are strongly correlated with cultural proximity. Figures 7 and 8 present binned scatterplots that residualize out cluster and district fixed effects.

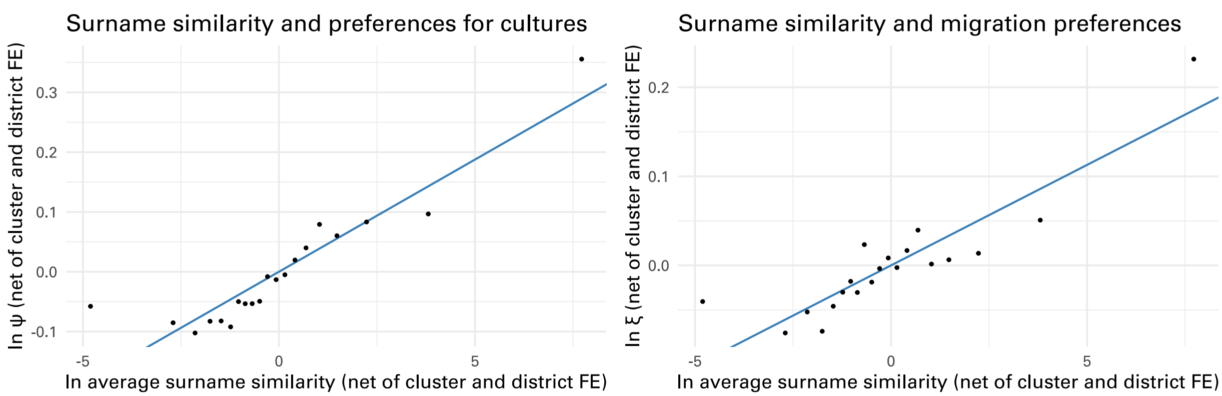

Figure 7: Binned scatterplots of ψ, ξ and the average similarity of surnames in a district to those in a given cluster in 1800

The parameter v_d^θ captures factors common to a location that induce migration. During the period we study, Britain was an industrializing economy. We would expect that locations experiencing structural transformation out of agriculture and into manufacturing would attract migrants. Figure 9 shows binned scatterplots of the relationship between the log shares in agriculture and manufacturing, and the log of migration incentives, $\theta \ln v_d$. There is a strong

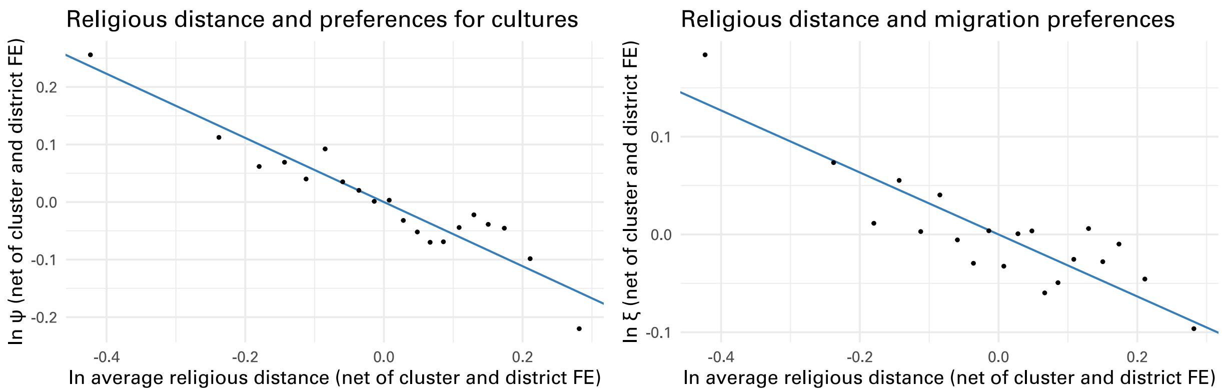

Figure 8: Binned scatterplots of ψ, ξ and the average Euclidean distance of religious denomination shares in a district to those in a given cluster in 1851

positive, though nonlinear, relationship between manufacturing and migration incentives, and a negative association between the share in agriculture and migration incentives.

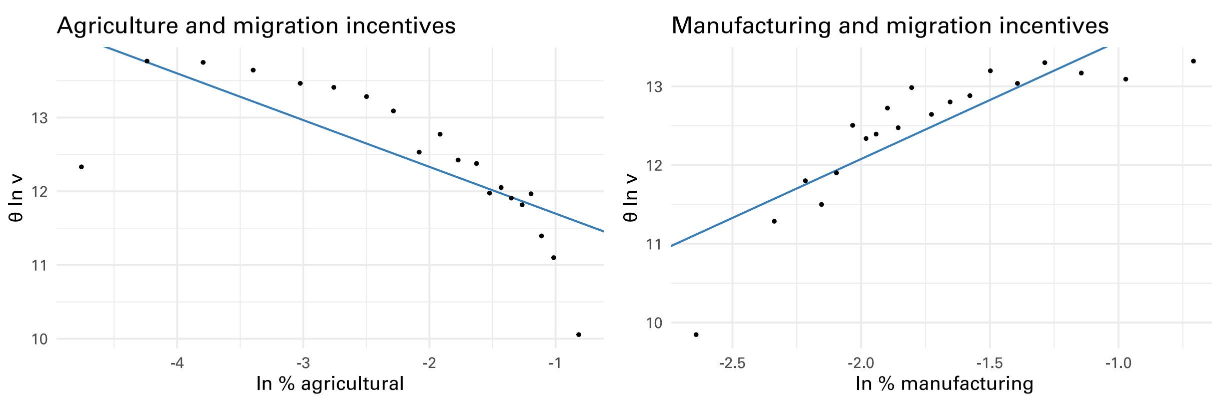

Figure 9: Binned scatterplots of log shares of employment in agriculture and manufacturing and $\ln v_d^\theta$

4.8 *Validating the Model's Predictions of Cultural Change in Response to Economic Change*

A counterfactual simulation in which we solve the model for 1851 economic parameters provides an additional way to test the model. Because we observe in 1911 people born between 1841 and 1860 in addition to the 1861–1895 generation that we use for our main analysis, we can examine how the changes in cultural choices between a counterfactual estimated using 1851 parameters and the observed values for the 1861–1895 cohort correlate with the difference between the cultural choices of the 1841–1860 and 1861–1895 cohorts.

These two differences should be positively correlated: the choices of the 1841–1860 generation were influenced by the distribution of economic activity around 1851. This correlation is however a strong test. We would not necessarily expect non-economic cultural preferences (ξ_d^k and ψ_o^k) to be stable over long periods. The large changes in the popularity of names associated with particular clusters are likely in part due to features of the data, and not due to the kind of cultural choices our model is designed to capture (for instance, while we do observe patterns in intermarriage that mirror those in names, the magnitude of the changes are considerably smaller). Note that this test is considerably stronger than, for instance, checking whether cultural choices for the 1841–1860 cohort correlate with those in the 1851 counterfactual, as we would expect both to be strongly correlated with cultural choices in the 1861–1895 cohort.

Economic fundamentals enter the model through three sets of exogenous variables. 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, which alters the starting population in different locations. Third, the costs of migration change over time.

As a first step towards quantifying the effects of economic change on cultural choices, we estimate the economic components of the model using data on district-to-district lifetime migration from the 1851 Census. We do not have data on naming, and so cannot analyze cultural choice in the same way we can using the 1911 data. We estimate migration pulls and costs using a gravity regression: we regress migration flows against log distance and an indicator that the origin district is the same as the destination district, with fixed effects for the origin and destination. Note that this method differs from how we estimate these parameters in the 1911 data as we lack data on names and thus on cultural groupings. We calculate migration pulls as the exponential of the destination fixed effects, raised to the power of 0.74. We rescale by that scalar because in a regression of $\theta \ln v_d$, estimated using the 1911 data, against the destination fixed effects from a gravity regression fit using only origin-destination flows in the 1911 data, the coefficient on the destination fixed effects is 0.74 (the R^2 is 0.98). This rescaling accounts for the fact that the model estimated without cultural groupings ignores that the attractiveness of a destination in our model increases with the number of members of a given culture living there. We calculate migration costs using the coefficients from the gravity regression. We also use the starting populations of migrants in 1851 to provide an estimate of starting populations.

Table 4 shows the results of regressions at the district-cluster level of the change in log σ between the two cohorts against the difference in log σ between the 1861–1895 cohort and a

	$\ln \sigma_{1861-1895} - \ln \sigma_{1841-1860}$			
	(1)	(2)	(3)	(4)
$\ln \sigma_{1861-1895} - \ln \sigma_{1851 \text{ counterfactual}}$	0.284	0.209	0.225	0.125
	(0.039)	(0.041)	(0.030)	(0.028)
District FE		x		x
Cluster FE			x	x
N	7345	7345	7345	7345
R^2	0.013	0.077	0.526	0.590

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 utilities 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 4: Relationship between change in log σ relative to 1851 counterfactual and relative to 1841–1860 cohort

counterfactual simulation using 1851 values of real wages, migration costs, and populations. There is a strong positive correlation between the two, 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-2 compares the spatial distribution of the percentage change in the share choosing the Southeast English cluster. While clearly not perfect, the model reproduces the broad spatial pattern of the observed change.

5 COUNTERFACTUALS

Given estimates of the structural parameters, we can solve for equilibrium cultural choices and migration flows under counterfactual scenarios. Doing so quantifies the effects of different changes to economic fundamentals on culture, taking into account general equilibrium effects. These simulations hold cultural preference terms, ξ_d^k and ψ_o^k constant. They thus provide estimates that isolate the mechanisms highlighted by the model from other cultural changes.

Before discussing the counterfactuals in detail, we note several features. Our estimates indicate that economic changes contributed to the weakening of regional cultures and the growth of the Southeast culture. This average result masks considerable heterogeneity across regions. Places that did not industrialize, but were close to places with distinct cultures that did industrialize, tended to lose their local culture. Places on the periphery that industrialized tended to retain their local culture. Changes to both the location of industry and migration costs contributed to the rise of the Southeast culture. We then use the model to estimate location-specific cultural responses to economic shocks, and find that economic development bolsters local culture on the periphery. Last, we flip the question and examine how total assimilation into one culture affects economic outcomes. Total assimilation into the Southeast culture would have increased economic output by diverting population to regions with higher wages.

5.1 *Estimates of the Effects of Economic Change on Culture*

We begin our counterfactual analysis with the aggregate effects of economic changes between 1851 and 1911 on the popularity of different cultures. Table 5 shows the changes in different quantities under counterfactual scenarios relative to the observed outcome. In the first three simulations we separately replace the migration attraction of all locations, starting populations, and migration costs, with their 1851 analogues. In the fourth we use 1851 estimates for all three variables. Altering starting populations should mechanically shift many of the variables: growing the population in locations with a strong taste for culture k should increase the number of people assigned to k. We therefore examine both the raw percentage changes in the number choosing different cultures and migrating, and a percentage change that fixes the starting populations at their 1911 values, and so just captures changes due to differences in choices of cultures.

Economic change increased the popularity of the Southeast cluster and decreased the popularity of regional cultures. Replacing 1911 values with their 1851 analogues decreased the popularity of the Southeast culture by 5%, and increased the popularity of home cultures by 1%. The rise of the Southeast is attributable to changes in real wages across locations; falling migration costs account for the decline of regional cultures. Replacing 1911 real wages with 1851 values results in a 1% decrease in the share choosing their home culture, and a 5% decrease in the share choosing the Southeast English culture. This pattern of results, which suggests that industrialization over the course of the 19th century increased both regionally-particular cultures and the hegemonic culture, is mirrored in the results from changing starting populations. These two effects appear contradictory, as we would expect the rise of the Southeast to come at the expense of regional cultures. Examining the spatial distribution of counterfactual cultural changes (Figure 11) reveals that reductions to the home culture in the counterfactuals are driven by the Southeast and middle of England, and by parts of South Wales and the Northeast of England that industrialized during this period. In both simulations, the Herfindahl-Hirschman Index of cultural shares decreases relative to the observed value in 1911, suggesting that economic change led to more local concentration in cultural choices, which is consistent with the growing size of the Southeast

Table 5: Counterfactual estimates

	Home culture pop.		S E Culture pop.		σ HHI	Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	%Δ	fixing pop.
1851 v_d^θ	-1.342	-1.342	-4.764	-4.764	-1.657	-3.625	-3.625
1851 starting populations	-5.719	-0.671	-9.730	-0.682	-0.502	3.546	0.067
1851 migration costs	2.688	2.688	-0.812	-0.812	1.542	-25.263	-25.263
All three	-1.961	1.123	-12.174	-4.803	0.992	-30.209	-30.231

This table shows percentage changes in the population choosing each district's home culture, the population choosing the Southeast English culture, the Herfindahl-Hirschman Index of culture shares, 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. The first row replaces destination utilities with their estimated 1851 value, the second replaces the populations migrating from each location with their 1851 value, the third replaces migration costs with those estimated from 1851 data, and the fourth replaces all three.

cluster. Changes to migration costs, which led to large decreases in the share migrating outside the district of birth relative to 1911, pushed in the opposite direction in the aggregate. In the counterfactual, the home culture population increases, the Southeast culture population decreases, and concentration increases. While the net direction of these results differs from the other two, the spatial patterns are broadly similar. The increase in home cultures is due to a much larger increase in the share choosing home cultures in the periphery.

While the net effect of economic change was to diminish regional cultures and increase the Southeast English culture, the effects were extremely heterogeneous across regions. Figure 10 shows kernel densities of the changes in shares choosing the home and Southeast cultures under these different counterfactuals. Across both outcome variables and all counterfactual scenarios, for a large share of districts the sign of the effect was in the opposite direction to the average.

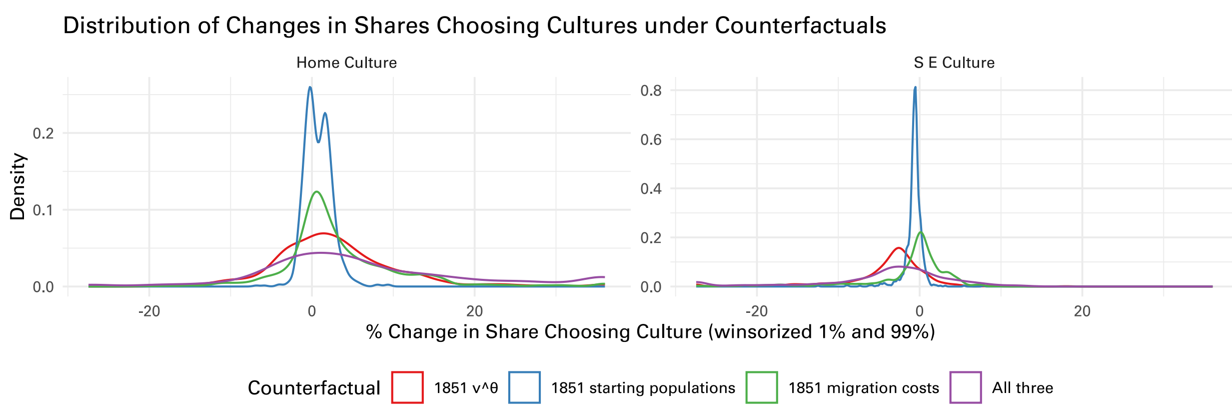

Figure 10: Densities of percentage changes in shares choosing home culture and Southeast England culture under different counterfactual scenarios

The spatial distribution of industrialization in this period partially accounts for this heterogeneity. The largest counterfactual increases in the share choosing the home culture are in parts of North Wales, Northern England, and East Anglia (the far east of England) that did not industrialize during this period but were close to other regions that did (see Figure 2). In these places, industrialization in other clusters—South Wales and the areas around Newcastle and London—increased the incentive to assimilate into these proximate clusters. The uneven pattern of industrialization also explains heterogeneity in the effects on the popularity of the Southeast English culture. The regions in which this culture increases in the counterfactuals are the Southwest of England and the South Wales coalfield. Industrialization in South Wales pulled migrants from Southwest England who would have otherwise gone to the Southeast, while increased opportunities in the Southeast decreased the rate of migration from the Southeast to the Southwest.

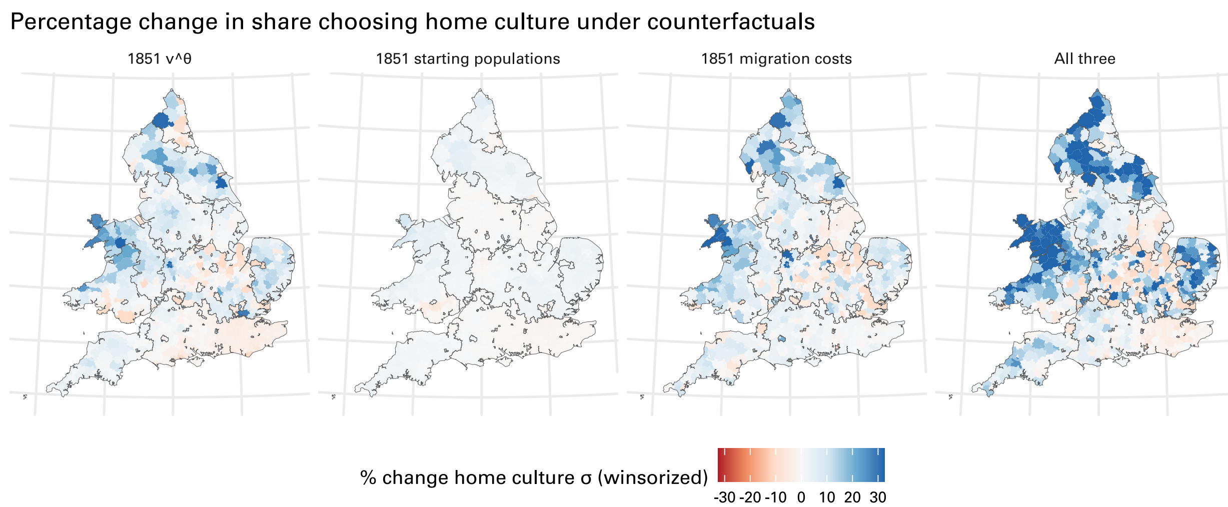

Figure 11: Spatial distribution of percentage changes in shares choosing home culture under different counterfactual scenarios

In addition to changes in industrialization, the interaction between changing migration costs and the level of industrialization contributed to the patterns we observe. The simulations which just change migration costs display a similar pattern to those that change the spatial distribution of economic activity. Falling migration costs also lead to decreases in the share choosing the home culture in places near industrial hubs allocated to different clusters. Reducing the cost of migration increases the probability of migration outside the cluster. Similarly, the counterfactual decreases in the share choosing the Southeast cluster are largest on the periphery, in Wales and North England. Decreasing the cost of migration increased

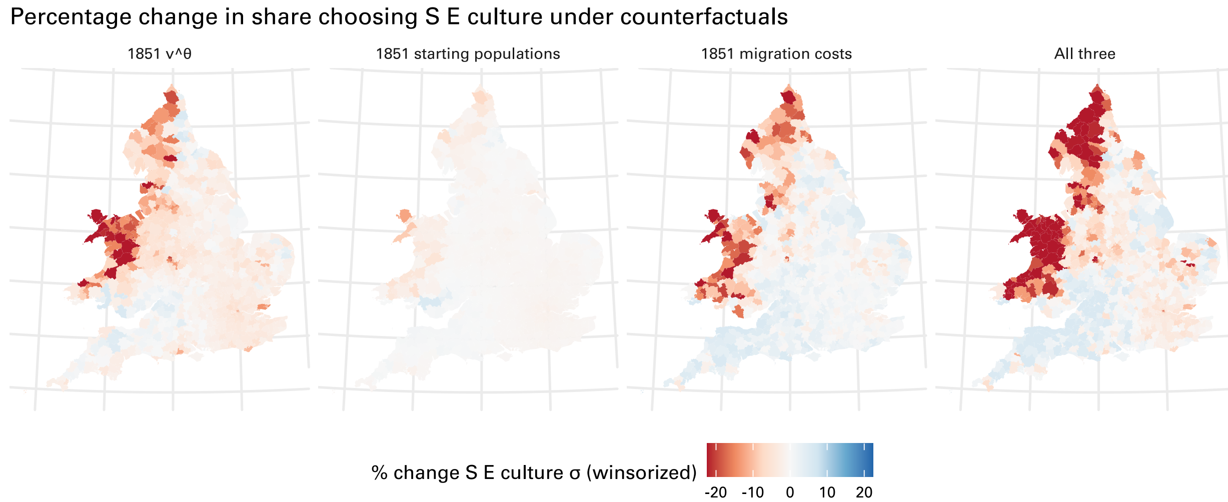

Figure 12: Spatial distribution of percentage changes in shares choosing Southeast culture under different counterfactual scenarios

the pull from the periphery to London, but did not pull in other migrants, who would have brought other cultures, to the periphery.

5.2 *Model-Implied Elasticities of Cultural Choice to the Real Wage*

Section 3 raised the question of how cultural choices evolve in response to industrialization. We provide a model-based answer—and further quantify heterogeneity—by examining how cultural choices in each location change in response to development in that location, holding all other parameters constant. For each district, we run a separate simulation in which we increase v_d^θ by 1% in that district, and record the change in cultural choices in that district. The answer is that it depends upon geography. As shown in Figure 13, 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.

Table 6 examines the observed relationship between changes in v_d^θ, changes in industrialization, and changes in cultural choices. In models (1) and (4), we regress the change in log σ at the district-culture level between the 1841–1860 and 1861–1895 cohorts against the changes in the log of v_d^θ and log manufacturing workers per capita, between 1851 and 1911. We interact the independent variable with the share choosing the culture in the early cohort.[8] The rationale for this specification is that if industrialization bolstered local cultures,

[8]Note that since choices of different cultures are by construction negatively correlated within a district,

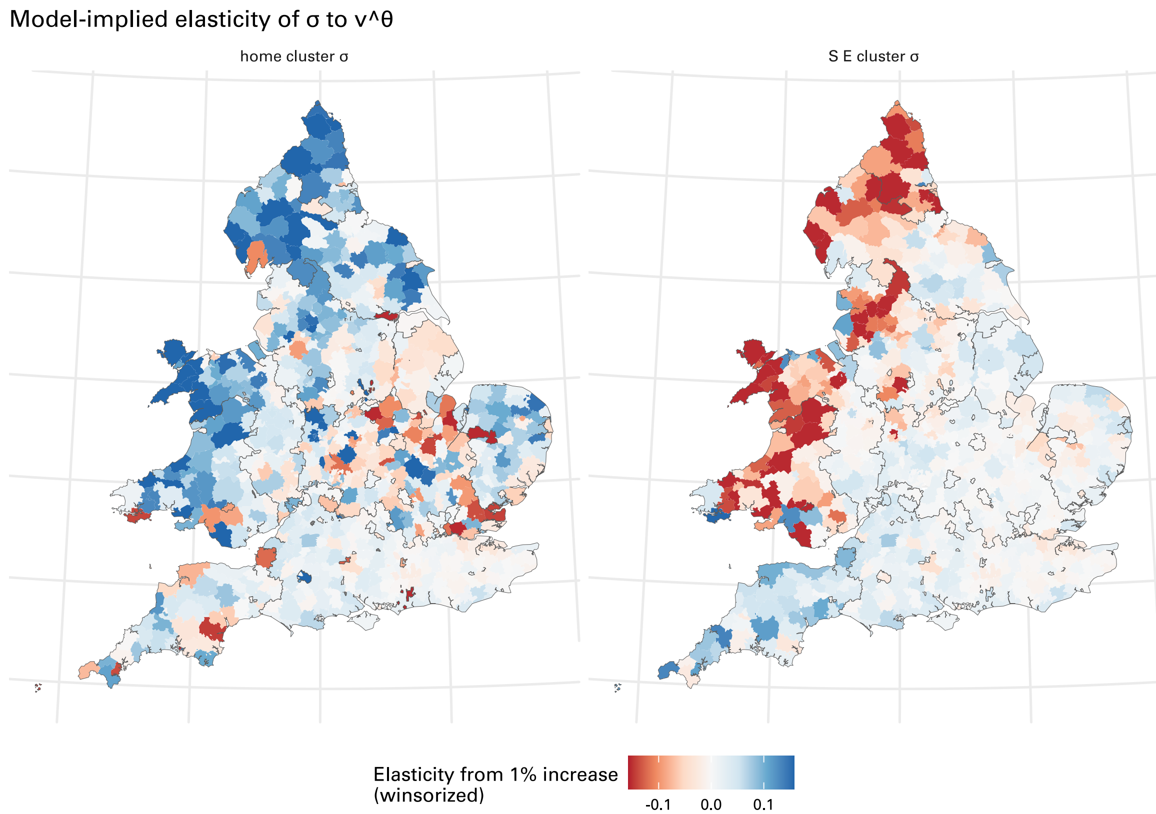

Figure 13: Elasticities of cultural choices to migration pull factors

We run one counterfactual simulation for each district in which we increase the migration pull 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 cluster as allocated based on surnames; the map on the right shows the elasticity of σ for the Southeast cluster.

it should increase the share choosing cultures which were popular in a period of reduced industrialization, and so the interaction term would be positive, while if it undermined them, the interaction term should be negative. We do not find evidence of a clear effect in either direction. In models (2) and (5) we instead interact industrialization with our model-implied elasticity. We find that the interaction is positive, precisely estimated, and robust to the addition of district fixed effects. This exercise demonstrates the utility of our structural approach relative to a reduced-form analysis. We need theoretically-motivated measures to make sense of heterogeneity.

What explains 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.

and industrialization varies at the district level, the coefficient on industrialization in an un-interacted model should be close to zero.

	$\Delta \ln \sigma$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\sigma_{1841-1860}$	-1.886			-1.787		
	(0.107)			(0.086)		
Model-implied elasticity		0.003	0.004		-0.004	-0.002
		(0.001)	(0.001)		(0.001)	(0.001)
$\Delta\,\theta \ln v_d$	0.006	0.006				
	(0.009)	(0.005)				
$\Delta\,\theta \ln v_d \times \sigma_{1841-1860}$	-0.020					
	(0.057)					
$\Delta\,\theta \ln v_d \times$ elasticity		0.005	0.004			
		(0.001)	(0.001)			
Δ ln Mf workers				0.009	0.004	
				(0.008)	(0.004)	
Δ ln Mf workers $\times\, \sigma_{1841-1860}$				-0.060		
				(0.044)		
Δ ln Mf workers $\times$ elasticity					0.003	0.002
					(0.001)	(0.001)
District FE			x			x
N	7345	7345	7345	7300	7300	7300
R^2	0.562	0.524	0.590	0.562	0.525	0.592

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. Models (1) and (4) examine the direct effects of changes in migration incentives or manufacturing on cultural choices interacted with the share choosing the culture in the 1841–1860 cohort, (2), (3), (5) and (6) interact these economic fundamentals 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. All models include cluster fixed effects, (3) and (6) add district fixed effects. Standard errors clustered by district in parentheses.

Table 6: Relationship between industrialization and cultural change, moderated by model-implied elasticities

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 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 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 4. Conditional on district fixed effects, this measure is very strongly correlated with the general equilibrium elasticities (Table A-4), with an R^2 of 0.98.

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. A positive shock to the real wage in location o influences people's choices for culture k more than other cultures if there are non-economic benefits to being of culture k if one lives in o. These components explain why for most locations the general equilibrium elasticity for the home culture is positive. The denominator captures migration opportunities across locations for an individual of culture k in location o. The magnitude of in-migration of people of culture k (which increases m_o^k in the numerator) relative to opportunities to migrate to other locations as a member of culture k (the denominator) accounts for the core-periphery pattern we observe in Figure 13. In locations on the periphery, the migration pull of the Southeast cluster is small, but the in-migration of members of the Southeast culture is smaller, which makes the appeal of the Southeast for those remaining in the district weaker.

Table 7: Counterfactual estimates from homogenizing culture

Counterfactual	%Δ Ave. v_d^θ	%Δ share migrant
Removing non-S E ψ_o^k	2.447	1.241
Removing non-S E ψ_o^k and fixing $\xi_d^k = 1$	-0.394	0.136
Doubling S E ψ_o^k	0.832	0.278
Halving S E ψ_o^k	-0.553	-0.117
Fixing ψ_o^k to have same average for all cultures	-0.489	0.094

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.

5.3 *Effects of Cultural Homogeneity on Migration and Economic Outcomes*

We now turn to 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 7 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 2.4%. 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 7 corresponds to around a 1.6% increase in average wages. Morten and Oliveira (2023) estimate a migration elasticity of 4.5 in Brazil, corresponding to a 0.5% 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, 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 particular preferences of members of the Southeast culture for particular 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 14 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.

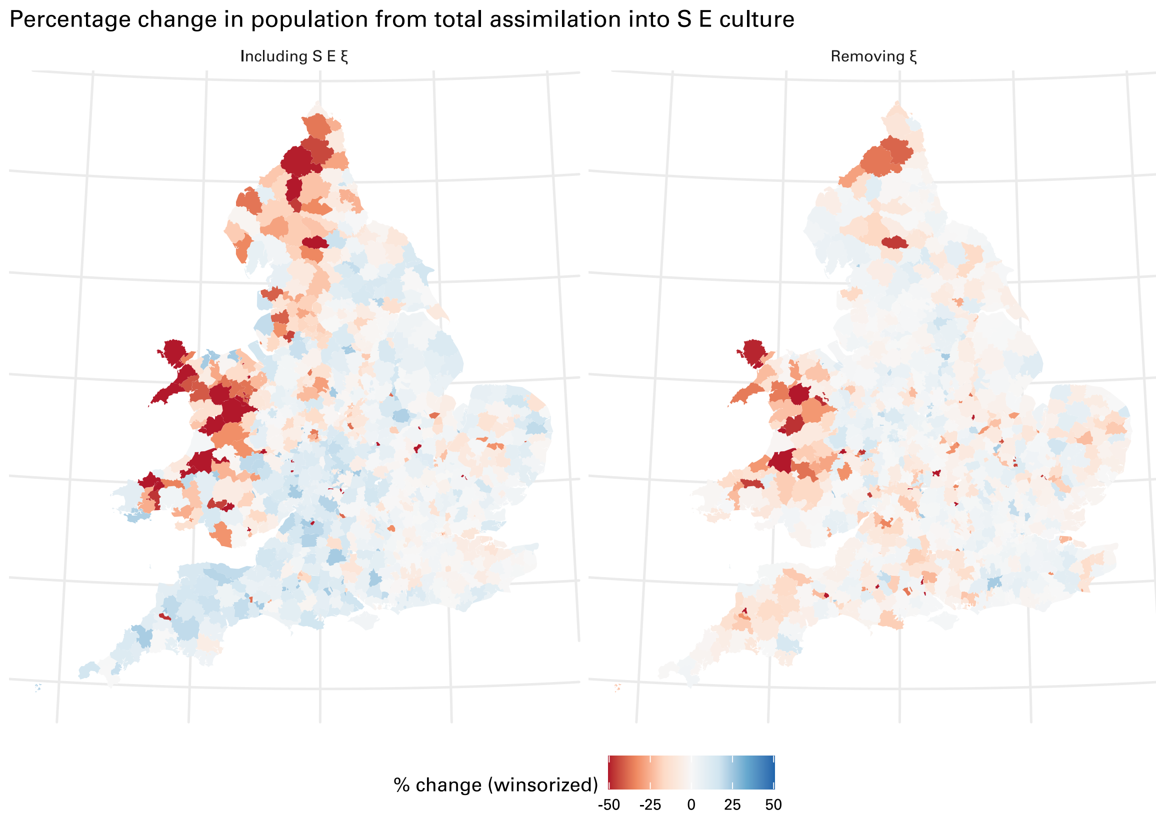

Figure 14: Changes in population distribution from homogenizing culture

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.

6 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 loss of local identities and a shift towards the culture of London, but also heterogeneity in the degree of cultural change. We develop and estimate a quantitative spatial model that allows us to characterize this heterogeneity. The model indicates that industrialization leads to the dilution of local cultures in central locations, which receive many migrants from different cultures, but promotes cultural retention in peripheral areas by lowering incentives to out-migrate. Our model further allows us to quantify the contribution of cultural homophily in migration decisions and its economic implications. Preference for local cultures moves people away from locations that could offer them higher wages, lowering average income.

Our findings both confirm and expand insights of modernization scholars on the homogenizing effects of industrialization and the role of labor migration (Deutsch, 1969; Gellner, 2006; Weber, 1976). The conventional wisdom in this literature is that migration drives the adoption of common identities and the loss of cultural distinctiveness. While we find this to be true on average, we also show that the effects of migration differ depending on the context. The modernization literature has identified multiple mechanisms through which industrialization led to both homogenization and differentiation of cultures and identities (Deutsch, 1969; Gellner, 2006). Our results suggest that labor migration is a distinct, but underappreciated channel that likely also contributed to both these developments.

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 cases of Catalonia and the Basque country in Spain, or Bavaria in Germany also correspond to 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.

More broadly, the migration link between industrialization and cultural change is relevant for the 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). Our findings in a context in which the state did not engage in any of these strategies suggest not only that spatial changes in the pattern of economic activity can be sufficient forces for cultural unification, but also that industrial policy, besides its more well-documented role in restructuring economic activity (Juhász, Lane and Rodrik, Forthcoming), can be used as a tool for 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.

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.

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.

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.

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.

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. Publisher: The University of Chicago 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'.

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

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

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.

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. *Nations and Nationalism.* Oxford: Blackwell Publishing.

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

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.

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

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

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.

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

Marx, Karl and Friedrich Engels. 1978. 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.

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.

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.".

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*

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.

Online Appendix

List of Tables

List of Figures

	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 A-1: People in 1851 tended to migrate within-cluster

A additional tables and figures

	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 A-2: People in 1851 tended to marry within-cluster

	Δ ln share home culture		
	(1)	(2)	(3)
Δ ln out-migrants / pop	−0.281		
	(0.061)		
Δ ln in-migrants / pop		−0.066	
		(0.026)	
Δ ln Mf workers			−0.017
			(0.031)
N	783	782	783
R^2	0.062	0.027	0.002

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. In (3), the change in the log number of manufacturing workers, 1851–1901. Observations are weighted by the number allocated namescores in the 1851–1860 cohort. Robust standard errors in parentheses.

Table A-3: Relationship between migration and the popularity of the home culture

	General equilibrium elasticity			
	(1)	(2)	(3)	(4)
Partial equilibrium elasticity	4.160	214.367	4.116	214.455
	(0.540)	(1.596)	(0.537)	(1.591)
District FE		x		x
Cluster FE			x	x
N	7345	7345	7345	7345
R^2	0.011	0.977	0.017	0.978

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 A-4: Relationship between partial and general equilibrium elasticities

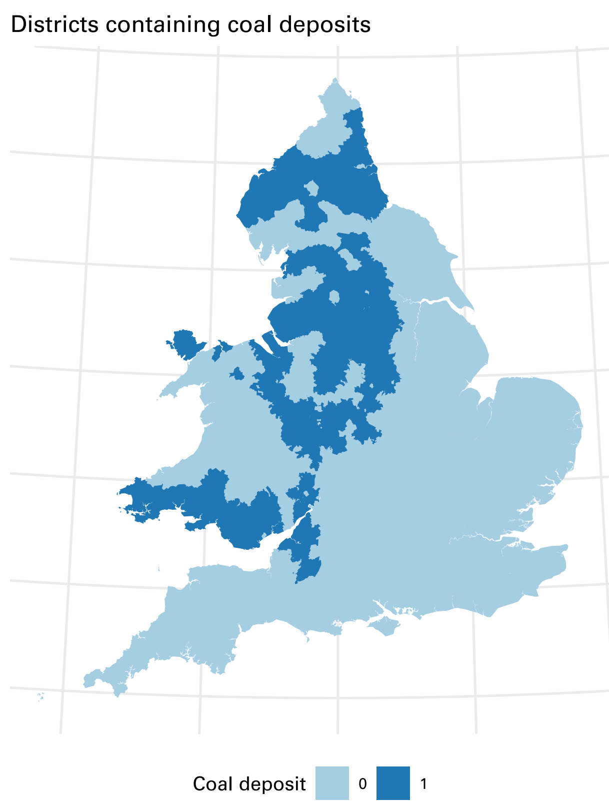

Figure A-1: Registration districts containing coal deposits

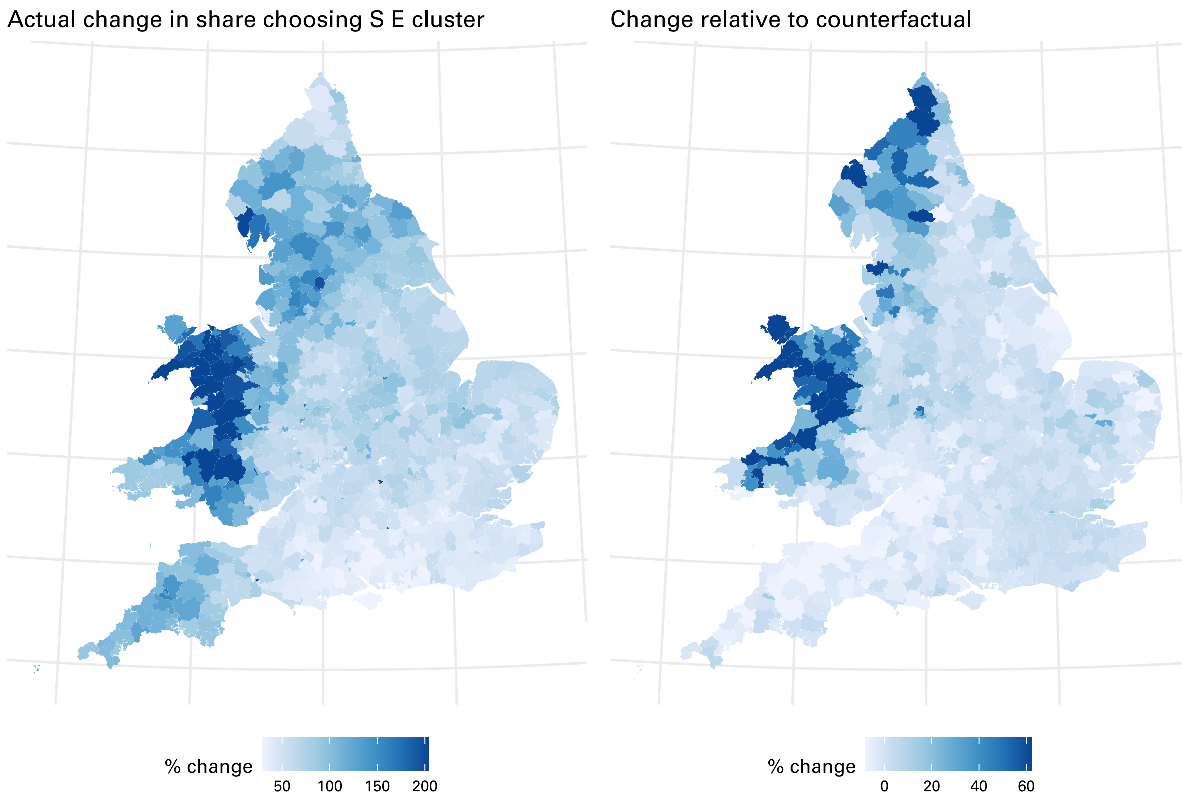

Figure A-2: Change in log share choosing the Southeast England cluster. Left panel is observed change between those born 1841–1860 and those born 1861–1895, right panel is change between counterfactual using 1851 estimates of migration costs, population, and destination utilities, and observed values for those born 1861–1895

B Quantifying changes in migration costs

This appendix discusses the estimation of migration costs in a gravity model, and how this approach allows us to distinguish between the effects on migration of changes to the incentives to migrate to different destinations and changes to costs. Suppose that people migrate to locations based in part on differing utilities in the destination v_d, in part based on inverse migration costs δ_{od}, and in part based on idiosyncratic preferences for specific locations ε_{di} drawn iid from a Fréchet distribution with shape parameter θ. This setup is similar to the model of migration in Section 4, but differs in ignoring culture and homophily. The utility individual i receives from migrating from o to d is thus:

$$u_{odi} = v_d \delta_{od} \varepsilon_{di}$$

From the properties of the Fréchet distribution, the number of people choosing to migrate from location o to location d is

$$m_{od} = \frac{(v_d \delta_{od})^{\theta}}{\sum_{j=1}^{N} (v_j \delta_{oj})^{\theta}} n_o$$

where n_o is the number of individuals originating in o. Taking logarithms gives the following linear equation:

$$\ln(m_{od}) = \underbrace{\ln\left(v_d^{\theta}\right)}_{\text{Destination FE}} + \underbrace{\theta \ln(\delta_{od})}_{\text{ln migration cost}} + \underbrace{\ln\left(n_o / \sum_{j=1}^{N} (v_j \delta_{oj})^{\theta}\right)}_{\text{Origin FE}}$$

If we parameterize migration costs as consisting of geographical distance and a binary effect of leaving one's location, $\delta_{od}^{\theta} = \text{distance}_{od}^{\beta_1} \exp\left(\beta_2 \mathbf{1}_{\{o=d\}}\right)$, we can estimate the following model:

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

The destination fixed effects γ_d provide an estimate of incentives to migrate to destinations, the coefficients β_1, β_2 on log distance and the indicator that the origin and destination are the same provide a measure of migration costs.

Given these estimates, and the population at each origin, we can predict the number of migrants from each origin to each destination. We can also decompose the contribution of different components to changes in migration rates over time, for instance by fixing origin populations and migration costs and varying the destination fixed effects.

Figure A-3 shows predicted rates of out of cluster migration from gravity models. The red line uses year-specific population and destination utilities and migration costs estimated from year-specific data. The green and blue lines use year-specific estimated migration costs and destination utilities, respectively, and fix the other variables at their 1851 levels. This figure shows that while changes in destination utilities contributes to an increase in out of cluster migration 1851–1911, most of the increase in migration is attributable to costs.

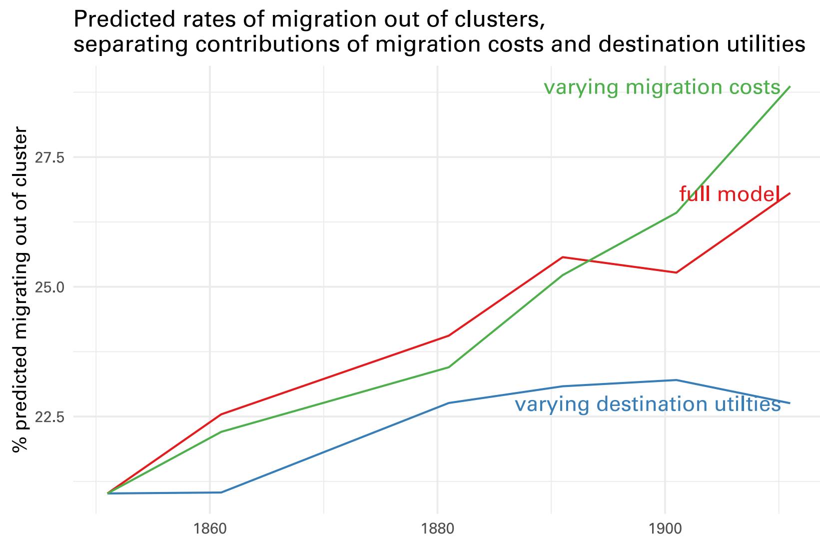

Figure A-3: Decomposing the change in inter-cluster migration into contributions from changing migration costs and destination utilities

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_{ij} 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. Remark 1 gives an additional result for cases where f_{ijh} takes as arguments equilibrium outcomes pertaining to other agents or locations as well. 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 $\Omega_o^k = \sum_{d=1}^N g_{od}^k(\mathbf{q}, \mathbf{\Omega})$ and $q_d^k = \sum_{o=1}^N h_{od}^k(\mathbf{q}, \mathbf{\Omega})$. In relation to Allen, Arkolakis and Li (2020), here we treat $N \times K$ location-by-culture pairs as the set of locations and Ω_o^k and q_d^k as the two types of equilibrium outcomes.[9] Note that the g_{od}^k function depends only on q_j^k, and is thus consistent with equation (1) of Allen, Arkolakis and Li (2020). The h_{od}^k function depends on both Ω_o^k and $\Omega_o^l \, \forall l \neq k$. 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

[9]One can think of the summations for Ω_o^k and q_d^k as summing over all location-culture pairs, but with the g and h functions returning zero if the culture does not equal k.

uniform bound on the sum of the elasticities.

The relevant matrix of uniform bounds on the elasticities is then

$$\mathbf{A} = \begin{pmatrix} \sup_{o,d,k} \left| \frac{\partial \ln h_{od}^k}{\partial \ln q_o^k} \right| & \sup_{o,d,k} \sum_{l=1}^{K} \left| \frac{\partial \ln h_{od}^k}{\partial \ln \Omega_o^l} \right| \\ \sup_{o,d,k} \left| \frac{\partial \ln g_{od}^k}{\partial \ln q_d^k} \right| & \sup_{o,d,k} \left| \frac{\partial \ln g_{od}^k}{\partial \ln \Omega_d^k} \right| \end{pmatrix}$$

Note that

$$\frac{\partial \ln h_{od}^k}{\partial \ln \Omega_o^k} = \varphi - 1 - \sigma_o^k \varphi, \; \frac{\partial \ln h_{od}^k}{\partial \ln \Omega_o^l} = -\sigma_o^l \varphi$$

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

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

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

$$\sum_{l=1}^{K} \left| \frac{\partial \ln h_{od}^k}{\partial \ln \Omega_o^l} \right| = \left| \varphi \left(1 - \sigma_o^k\right) - 1 \right| + \varphi \left(1 - \sigma_o^k\right)$$

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

For the other entries in $\mathbf{A}$,

$$\frac{\partial \ln g_{oj}^k}{\partial \ln q_j^k} = \frac{\alpha}{1 - \alpha}$$

and the two diagonals are zero.

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 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 A-4 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 A-4 shows these clusters.

The pattern we observe in Figure 3, 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 A-5 and A-6 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 4. 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

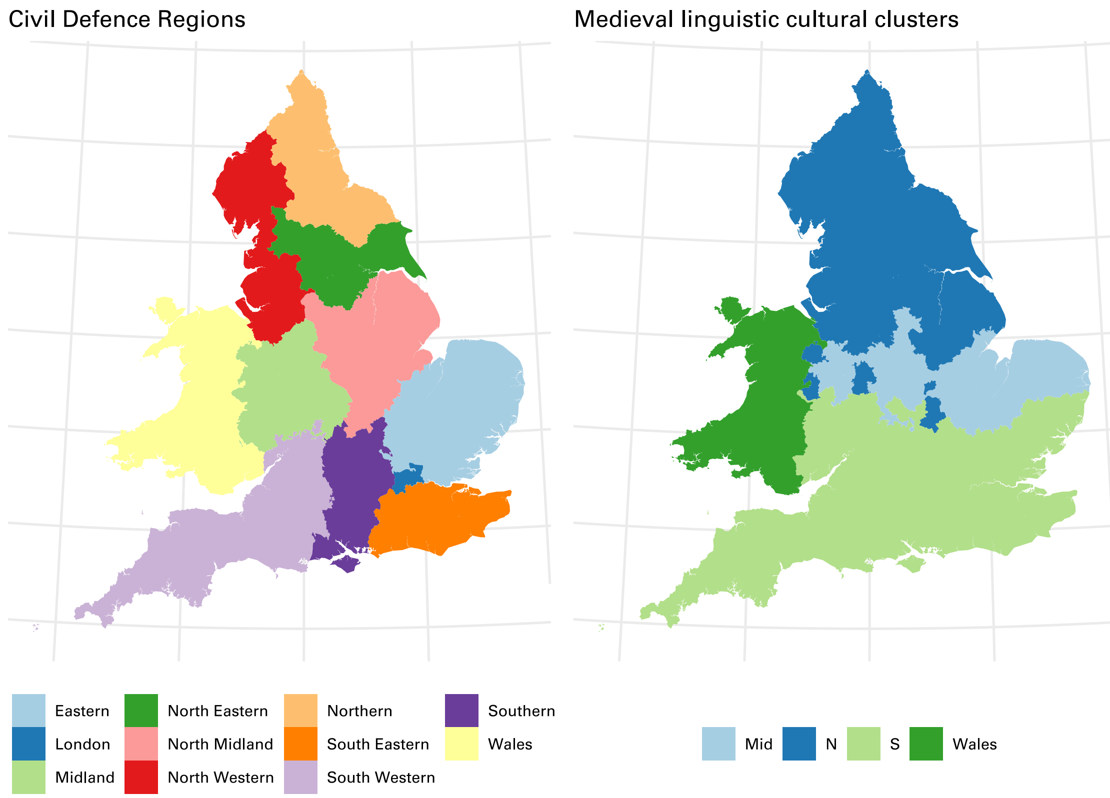

Figure A-4: Alternative cultural clusters

cluster in the case of the linguistic clusters).

Broadly speaking, our results with these alternative clusters are similar to those using clusters based on historical surnames. Tables A-5 and A-6 show the reduced form analyses with these new clusters, and Tables A-7 and A-8 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 A-6). 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 A-9 and A-10 show the same counterfactual outputs as Table 5, using the alternative clusters. The effects on the popularity of the Southeastern, Southern, and London cultures, are

	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 A-5: 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 A-6: Relationship between migration and naming patterns, with linguistic clusters

	$\ln\sigma_{1861-1895} - \ln\sigma_{1841-1860}$			
	(1)	(2)	(3)	(4)
$\ln\sigma_{1861-1895} - \ln\sigma_{1851\text{ counterfactual}}$	0.363	0.278	0.252	0.103
	(0.036)	(0.043)	(0.027)	(0.026)
District FE		x		x
Cluster FE			x	x
N	8999	8999	8999	8999
R^2	0.019	0.065	0.640	0.688

This table shows OLS estimates at the district-by-cluster level, replicating Table 4 using civil defence regions. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination utilities 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 A-7: Relationship between change in log σ relative to 1851 counterfactual and relative to 1841–1860 cohort, with civil defence regions

	$\ln\sigma_{1861-1895} - \ln\sigma_{1841-1860}$			
	(1)	(2)	(3)	(4)
$\ln\sigma_{1861-1895} - \ln\sigma_{1851\text{ counterfactual}}$	0.437	0.382	0.094	−0.086
	(0.060)	(0.069)	(0.028)	(0.026)
District FE		x		x
Cluster FE			x	x
N	3306	3306	3306	3306
R^2	0.020	0.080	0.799	0.865

This table shows OLS estimates at the district-by-cluster level, replicating Table 4 using linguistic clusters. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination utilities 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 A-8: Relationship between change in log σ relative to 1851 counterfactual and relative to 1841–1860 cohort, with linguistic clusters

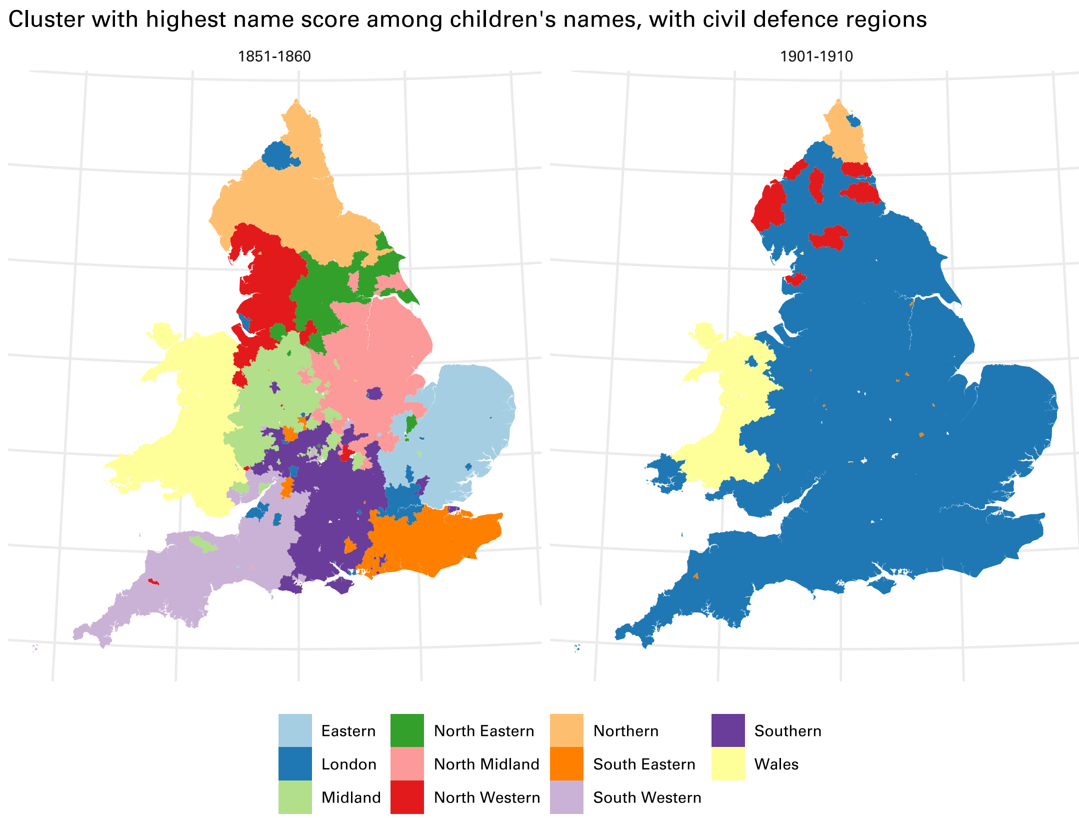

Figure A-5: Changing popularity of cultures, using civil defence regions as clusters

very similar across these specifications, which increases our confidence in our conclusion that 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 A-8 and A-10). Our estimates for the effects on the popularity of home cultures differ in their aggregate magnitudes but follow similar spatial patterns (Figures A-7 and Figures A-9). Using civil defence regions, the counterfactual simulations imply that economic changes had a larger negative effect on home cultures, which increase in popularity by 4% fixing population in the counterfactual, relative to an estimate of 1% given clusters based on surnames. This difference is likely attributable to the hegemonic culture, London, corresponding to a much smaller region. Areas in the Southeast outside London that are pulled towards London are coded in the main text as experiencing an increase in the popularity of the home culture, whereas with civil defence regions they are coded as experiencing a decrease. Differing boundaries affect our estimates using linguistic clusters in the opposite direction. Locations in the West of England that shifted towards the Southeast are coded as experiencing an increase in the home culture.

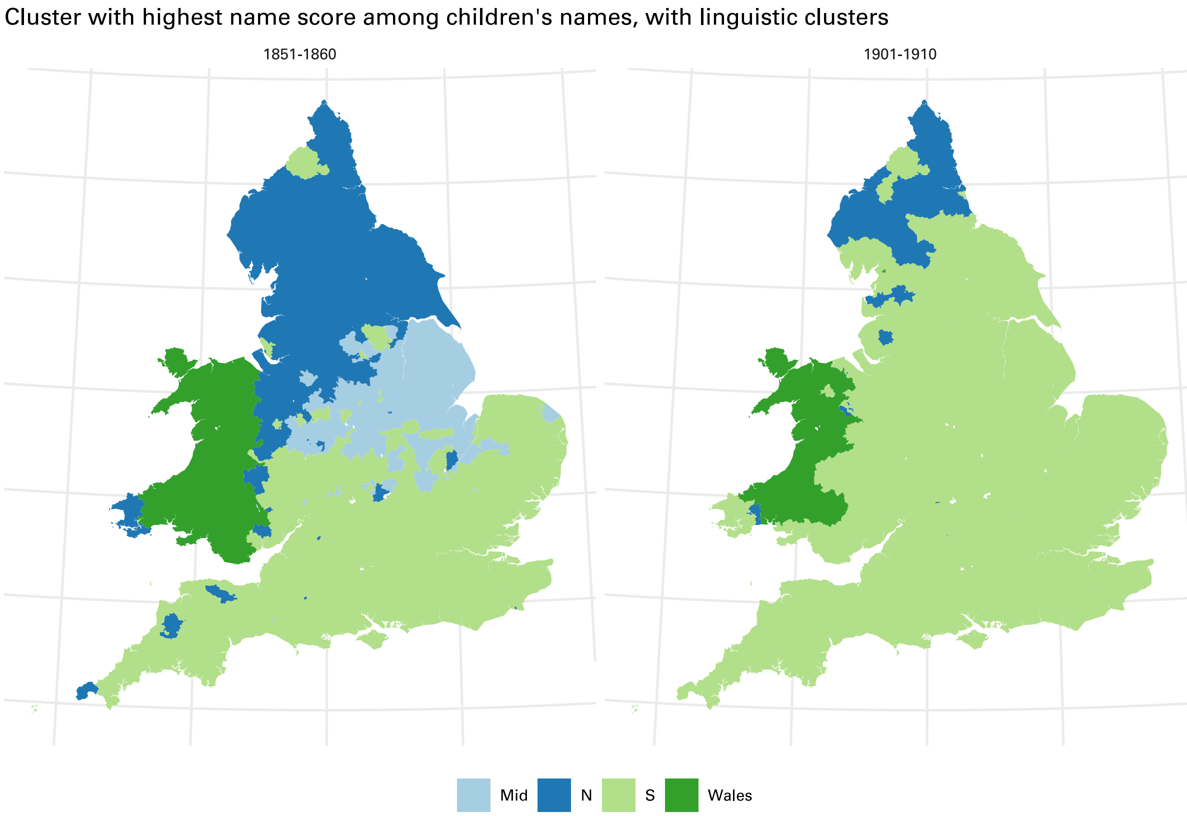

Figure A-6: Changing popularity of cultures, using linguistic clusters

Table A-9: Counterfactual estimates with civil defence regions

	Home culture pop.		London culture pop.		σ HHI	Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	%Δ	fixing pop.
1851 v_d^θ	0.837	0.837	-5.411	-5.411	-0.577	-3.540	-3.540
1851 starting populations	-14.267	-0.791	-10.759	-1.144	-0.552	3.508	0.065
1851 migration costs	4.281	4.281	-0.824	-0.824	1.722	-25.359	-25.359
All three	-8.276	4.024	-13.250	-5.577	2.185	-30.139	-30.162

This table shows the same counterfactual outputs as Table 5, using civil defence regions in place of clusters based on surnames before 1800.

Table A-10: Counterfactual estimates with linguistic clusters

	Home culture pop.		S culture pop.		σ HHI	Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	%Δ	fixing pop.
1851 v_d^θ	0.094	0.094	-3.384	-3.384	0.131	-3.893	-3.893
1851 starting populations	1.817	-0.426	-8.320	-0.256	-0.008	3.515	0.069
1851 migration costs	2.072	2.072	-0.499	-0.499	1.807	-24.941	-24.941
All three	5.008	1.773	-9.431	-3.130	3.140	-30.144	-30.219

This table shows the same counterfactual outputs as Table 5, using linguistic clusters in place of clusters based on surnames before 1800.

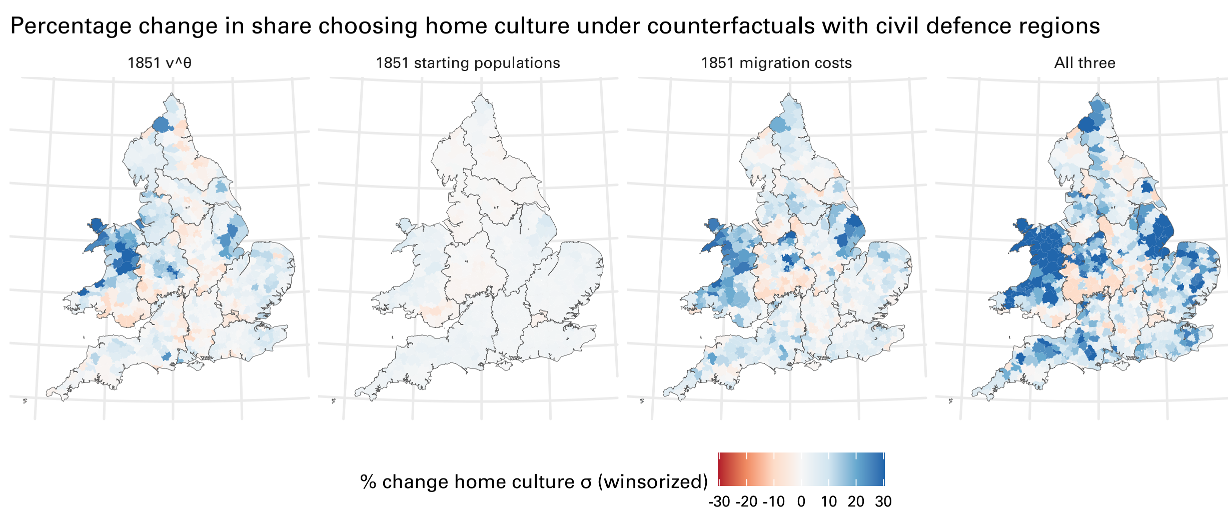

Figure A-7: Spatial distribution of percentage changes in shares choosing home culture under different counterfactual scenarios, using civil defence regions as cultural clusters

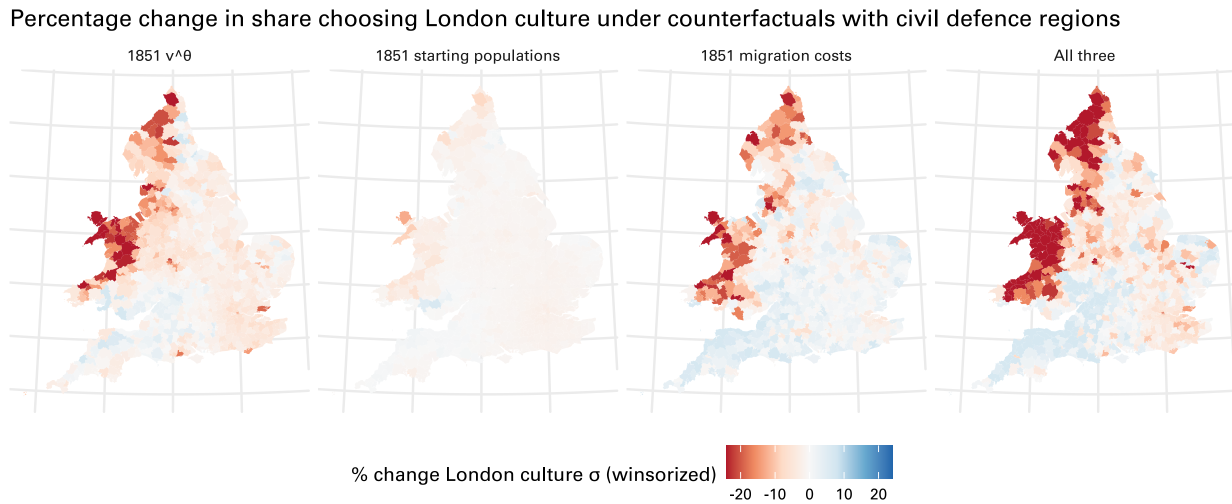

Figure A-8: Spatial distribution of percentage changes in shares choosing London culture under different counterfactual scenarios

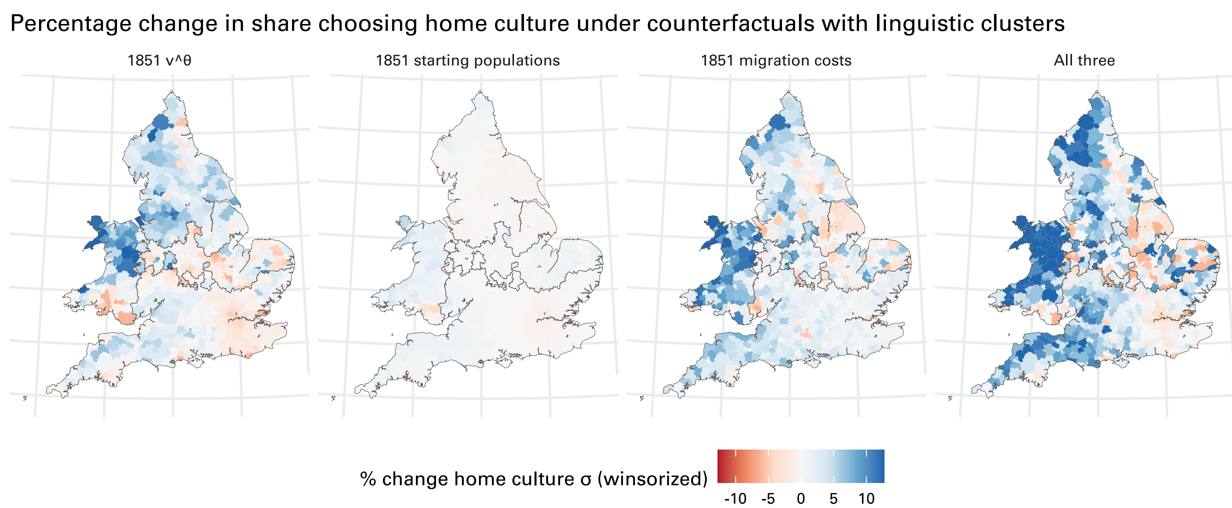

Figure A-9: Spatial distribution of percentage changes in shares choosing home culture under different counterfactual scenarios, using linguist clusters

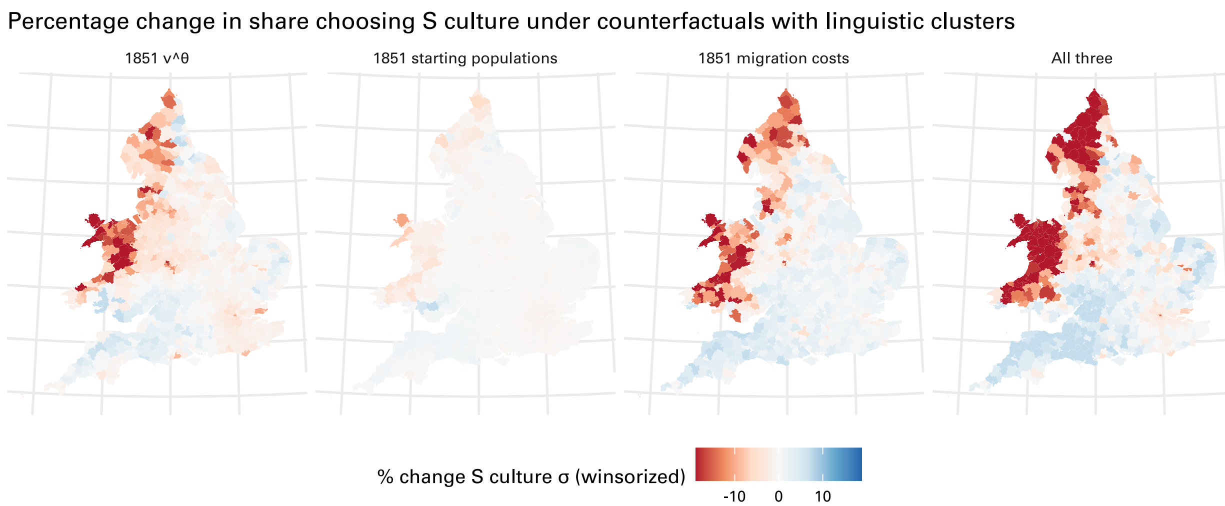

Figure A-10: Spatial distribution of percentage changes in shares choosing Southern culture under different counterfactual scenarios, using linguistic clusters