Migration and the Making of the English Middle Class*

Vasiliki Fouka† Theo Serlin‡

January 2026

Abstract

When do people identify with their class? Evidence from social psychology shows that individuals are more likely to identify with a group if they are similar to its members. We study early 20th century Britain and show that regional cultural heterogeneity combined with internal migration influenced class identity. We develop and validate a measure of class identity using naming decisions. Exploiting within-household variation, we show that migration patterns that increased the local share of culturally-distant workers reduced working class identification. Where migration increased the cultural distance of the working class, workers were less likely to join unions, voters were less likely to support the nascent Labour Party, and parliamentary candidates were less likely to target working class voters. By 1911, slower in-migration and rising local population growth reduced working class distance in urban areas, which also became strongholds of support for Labour. Migration alters social identity and creates political cleavages.

*We would like to thank Anna Grzymała-Busse, Elias Dinas, Felix Kersting, David Laitin, Biljana Meiske, Alain Schläpfer, Marco Tabellini and seminar participants at the 2025 ASREC Annual Meeting, ETH Zurich, IAST, EUI, the Berlin Workshop on the Political Economy of Identity, the Harvard Center for European Studies, King's College London, Berkeley Haas and the Paris School of Economics for useful comments and suggestions. Emil Almgren provided excellent research assistance.

†Stanford University. E-mail: vfouka@stanford.edu.

‡King's College London. E-mail: theo.serlin@kcl.ac.uk.

1 INTRODUCTION

When do people identify with their class? In the nineteenth century, political thinkers on both the right and the left assumed individuals would align with their class interests. Franchise extension was expected to lead to the triumph of the working class, socialism, and redistribution. Yet this future never materialized. And while class was central to electoral politics in much of twentieth-century Europe (Lipset and Rokkan, 1967), in various cases, especially the United States, class politics never consolidated (Przeworski and Sprague, 1986; Lipset and Marks, 2000).

This paper draws on social identity theory to understand class identification. Shayo (2009) formalizes a rich social psychology literature in a model in which individuals are more likely to identify with a group if they are more similar to its prototypical member. An unexplored implication from that framework is that migration which brings culturally distant members of the working class into an area will increase perceived distance to the working class among residents, and cause them to identify less with the working class. Group identification, in turn, drives policy preferences. If individuals identify less with the working class, they are less likely to support policies that benefit the working class. Industrialization at first retards the formation of class politics, because the process of transforming rural peasants into urban workers necessitates internal migration.

We explore this mechanism in the context of Britain at the turn of the twentieth century, a period of economic change and emergent class politics. We build on Shayo's (2009, 2020) framework of endogenous social identification to link class identity to the relative cultural proximity of different occupational classes. In this framework, (subjective) individual class identification depends on (objective) occupation and perceived distance from an occupational class. Individuals may identify as working class or move away from that identification toward complementary middle- and upper-class identities. Conditional on occupation, a greater perceived distance to the working class relative to the middle class—for instance, due to a higher presence of foreign workers than foreign professionals in the region of residence—weakens working-class identification and increases affinity with the middle class.

We then operationalize this theoretical link using census microdata from the full population of England and Wales between 1881 and 1911. We construct a measure of individual working class identification based on names given to children. The class valence of names is captured by their empirical distribution across occupations belonging to different classes, as defined at the time by administrative authorities. The Registrar-General distinguished between upper and middle class occupations—ranging from clerks to elite professionals such as bankers and lawyers—and working class occupations involving skilled or unskilled manual labor.

Distinctively working class names of children strongly negatively correlate with the father's occupational status. Among Members of Parliament, working class scores are higher for members of unions and Labour MPs and lower for MPs with elite education or members of elite social clubs.

We link this outcome to a measure of *relative class distance*—the extent to which an individual is culturally more distant from the working class than from the upper and middle class in their area of residence. Relative class distance varies within a parish across individuals of different origins. For an individual with a given birthplace, it increases when working-class residents in the parish are drawn from places that are culturally more distant from the individual's parish of birth. We compute cultural distance across parishes using historical intermarriage patterns, which strongly correlate with other dimensions of cultural distance, such as religion. With this measure in hand, we then exploit within-household variation across births and show that when families move to, or remain in, parishes where the working class is more distant to them, they become less likely to give their children working-class names, indicating a decline in identification with the working class. These results are consistent with parallel-trends assumptions in within-household naming behavior, are robust to alternative measures of name content and cultural proximity, and are not driven by changes in other features of names, such as regional distinctiveness.

Class identity translates into political behavior. We show this at the individual level using data on union membership. We link census microdata to two million individual records from fourteen large unions, collectively covering more than five thousand branches. We then compare individuals of the same origin and occupation across locations with different levels of relative class distance, holding constant propensities to unionize within residence–occupation groups. Higher relative distance from the working class decreases the likelihood of being a union member, a result driven by those in working class jobs.

These individual-level shifts scale up to electoral outcomes and elite responses. At the constituency level, in elections between 1892 and 1910, higher relative distance from the working class decreases Labour's vote share and the likelihood that the Party contests a seat, even after conditioning on rich covariates capturing the occupational structure of the electorate and various measures of migration. In our most restrictive specifications, a two–standard deviation increase in relative class distance corresponds to roughly a three–percentage–point decrease in the Labour vote share—a sizable effect given that Labour's average vote share during this period was below five percentage points. We also find that political candidates respond to shifts in the electorate's identification. Using text analysis of parliamentary election addresses, we construct a measure of working-class rhetoric that quantifies the extent to which candidates emphasize issues and language associated with working-class constituencies.

Candidates—including those within the same party—use less working-class rhetoric where cultural distance from the working class is higher.

Finally, we ask which migration patterns enabled mobilization in support of Labour by descriptively examining spatial heterogeneity in relative class distance. Between 1891 and 1911, urban areas experienced a decline in relative in-migration: as the share of locally born workers steadily grew by more than the population of migrants, the urban working class became more homogeneous. This aligns with the spatial distribution of the rise in Labour support, which was concentrated in cities. In rural areas, by contrast, rising out-migration of locally born workers increased relative distance to the working class and limited the formation of a working class identity. These patterns suggest an explanation for both the delayed timing of class-based political mobilization relative to the timing of industrialization, and its distribution across space.

This paper makes four contributions. First, we contribute to research on what shapes the strength of class as a cleavage, and in particular on the relationship between class and cross-cutting cleavages. Existing theoretical and empirical work shows that group divisions that cross-cut class weaken its relevance for political outcomes (Lipset and Rokkan, 1967; Roemer, 1998; Bartolini and Mair, 1990; Bartolini, 2000; Alesina, Glaeser and Sacerdote, 2004; Lee, Roemer and Van der Straeten, 2006).[1] Yet these models typically treat voter preferences and non-class cleavages as exogenous. We show instead how such cleavages can arise endogenously through a psychological mechanism of social identification, and we measure this process at the individual level using a novel, validated proxy based on first names.

Second, we contribute to the rapidly growing literature on the political effects of immigration. A wide body of evidence suggests that immigration tends to weaken support for redistribution (Luttmer, 2001; Abrajano and Hajnal, 2015; Alesina, Miano and Stantcheva, 2023) and increase the vote share of far-right parties (see Alesina and Tabellini for a review).[2] A related literature links ethnic and racial diversity to a similar set of outcomes (Alesina, Baqir and Easterly, 1999; Alesina and La Ferrara, 2000; Miguel and Gugerty, 2005) through mechanisms such as coordination failures (Habyarimana et al., 2007), preference heterogeneity (Alesina and La Ferrara, 2005) or outgroup bias (Gilens, 1999; Alesina and La Ferrara, 2000). We identify a new mechanism: immigration weakens psychological identification with the working class, as predicted by theoretical models of identity formation (Shayo, 2009, 2020).

[1] Research on the weakening of the class cleavage in other contexts highlights a distinct set of mechanisms that do not apply to our case, including Christian-Democratic parties (Kalyvas, 1996), the rise of the knowledge economy (Kitschelt, 1994), and the emergence of post-materialist values (Inglehart, 1977).

[2] There is considerable heterogeneity in this effect of immigration, mediated by mechanisms such as intergroup contact (Lonsky, 2021; Steinmayr, 2021) and benefits from migration to local communities (Gamalerio et al., 2023; Zhou, Grossman and Ge, 2023).

Moreover, we demonstrate this channel in a historical context—Britain at the turn of the twentieth century—that predates the mass immigration of the late twentieth century and was characterized by greater homogeneity and rising class consciousness. That we find such effects under these conditions suggests that our estimates may represent a lower bound on the potential impact of more culturally distant migration flows.

Third, we provide empirical evidence on perceived group distance as a determinant of identity formation. The idea that similarity to a prototypical group member shapes identification dates back to social identity theory (Tajfel and Turner, 1986) and has been formalized in later work (Shayo, 2009, 2020). While research has demonstrated the importance of perceived similarity for outgroup classification and attitudes (Fouka, Mazumder and Tabellini, 2022; Fouka and Tabellini, 2022; Hessami and Schirner, 2024), we lack evidence from outside the laboratory on how group composition affects which groups individuals identify with. We provide such evidence, showing that perceived group distance shapes identity choices in real-world settings.

Fourth, we contribute to historical debates on why the labor movement in Britain did not produce a proletarian revolution and why Labour's electoral success lagged expectations in the late nineteenth century (Hobsbawm, 1954; Pelling, 1965; Stedman Jones, 1983). Successive franchise extensions should have translated into greater working-class political power, a pattern likewise predicted by canonical political economy models (Meltzer and Richard, 1981; Acemoglu and Robinson, 2000, 2006). Prominent explanations have emphasized rising living standards during the mid-Victorian boom and the emergence of a "labour aristocracy" of high-paid skilled workers who developed reformist political positions (Hobsbawm, 1954; Cannadine, 1998). Yet other perspectives point out that the working class was not homogeneous, but was cross-cut by other solidarities, such as locality (McKibbin, 1990). We offer evidence for a specific version of this argument: that large internal migration flows driven by industrial development increased heterogeneity and slowed the consolidation of working-class identification.[3] With the exception of a few qualitative accounts (Pelling, 1967, p.42–46), this channel has received little attention in existing scholarship. We also provide an explanation for the time lag between the growth of the working-class population and its coalescence as a political bloc. Mass internal migration undermined class identity, a constraint that relaxed over time as descendants of earlier waves of urban migrants came to outnumber continuous new arrivals.

[3]Fresh (2024) shows that these migration flows, in the century preceding the period we study, also contributed to the growth of electoral contestation, a precondition for our findings.

2 HISTORICAL BACKGROUND

2.1 *Class identity and class politics, 1870–1914*

The last third of the nineteenth century saw British politics reorganize from divisions over religion, empire, and tariffs to cleavages centered on class. During the Second Industrial Revolution (1870–1910), industrial development accelerated, driven by advances in steel, chemicals, and engineering. Employment in manufacturing expanded sharply, so that by the first two decades of the twentieth century manual workers made up roughly 78 percent of the labor force, making Britain "one of the most working-class countries in the world" (McKibbin, 1998, p.106).

This social transformation coincided with successive extensions of the franchise. The Representation of the People Act of 1884 (the Third Reform Act) extended the concessions of 1867 from the boroughs to the countryside, enfranchising men paying an annual rental of £10 or holding land of equivalent value. Although regional variation in suffrage rights persisted, the Act greatly expanded political participation among the growing working class. Examining electoral registers in the working class East End of London, Brodie (2004, p. 52) concluded that the "overall occupational structure does not differ vastly from what would have been expected from a fully inclusive franchise," though wealthier members of the working class were more likely to be enfranchised.

By the 1880s a distinct class identity was emerging. It was visible in everyday behaviors that constituted a distinct working-class "culture," from dress—the characteristic worker's cap—to leisure activities like football, betting and gambling, pubs, and music halls (Stedman Jones, 1974; Hobsbawm, 1984*a*). Work itself stood at the center of this identity: English working men derived pride, self-worth, and social standing from their occupation (McKibbin, 1998). Labor historians emphasize that this identity permeated everyday life, reinforcing a sense of conflict between "us" and "them"—with "them" comprising the landed aristocracy, capitalists, and the lower middle classes. It rested on a moral code of solidarity, mutual aid, and cooperation (Hobsbawm, 1984*a*, p.190–191).

Trade unionism reinforced this sense of identity and became its main political expression. Union membership rose from about 750,000 in the 1880s (Adelman, 1996, p.10) to four million by 1914, aided by the legalization of strike action in the 1870s. Working-class political representation followed: first through "Lib-Lab" MPs allied with the Liberals, then through the Independent Labour Party (ILP) founded by Keir Hardie in 1893. In 1900 unions and socialist groups came together in the Labour Representation Committee (LRC), which, renamed as the Labour Party, entered Parliament in 1906 with 30 MPs. By 1914, voting

Labour had become the principal political expression of working-class identity.[4]

While class mobility was limited and most workers did not aspire to it (Hobsbawm, 1984*b*; McKibbin, 1998), the working class was far from homogeneous in its behaviors and identities. The sharpest divide was between the "respectable" and the "rough," the former embracing values often associated with the middle class—order, domesticity, sobriety, and thrift. Education and clerical employment provided a recognized path of advancement for those from working-class backgrounds, even though such careers were often less lucrative than skilled trades. Becoming a clerk has been described as being "lost to the working class forever" (McKibbin, 1998, 45). This internal gradient meant that individuals of similar origins retained some scope to negotiate their identities, aspirations, and behaviors.

2.2 *Internal migration*

Besides the rising importance of class, Britain's continuing industrialization spurred massive mobility. Throughout the 19th century, internal migration flowed from agricultural areas to cities; by 1851, more than 75 percent of the populations of northern industrial centers such as Manchester and Bradford had been born elsewhere (Feldman, 2000).

In the period that we study, 1881–1911, the most important migration-related development altering the cultural map at the local level was the continuation of rural to urban migration flows. Over time, those born in rural areas migrated over longer distances. Figure 1 shows that the share of people living outside their county of birth steadily increased between 1881 and 1911, particularly for agricultural locations.[5]

There was also significant migration from Ireland. By 1911, the Irish made up about 5 percent of the population of England and Wales (Cummins and Gráda, 2025). They settled overwhelmingly in northwestern centers such as Lancashire, Liverpool, and Manchester, where demand for unskilled labor in docks, textiles, construction, and mining was high.

Historians have emphasized that, in contrast to the uniformity of the middle class, workers remained strongly attached to regional and local identities and tastes (Hobsbawm, 1984*b*; McKibbin, 1998), which the arrival of outsiders could disrupt or threaten. A prominent example was the influx of Jewish migrants from Eastern Europe into London's East End, which provoked backlash among working class voters, pushing them toward the Conservative Party (Pelling, 1967). While this account emphasizes labor market competition, other mechanisms may also have weakened support for working class causes. In the next section, we outline a

[4]Bronner and Ziblatt (2018) show that class cleavages were already visible earlier, with the Liberal Party indistinguishable from Labour in its class appeals by 1914. The alignment of the working class with the Liberal left along class lines was already evident in early Victorian politics (Dewan, Meriläinen and Tukiainen, 2020).

[5]Figure A.1 shows a similar pattern using a continuous measure of distance to the parish of birth.

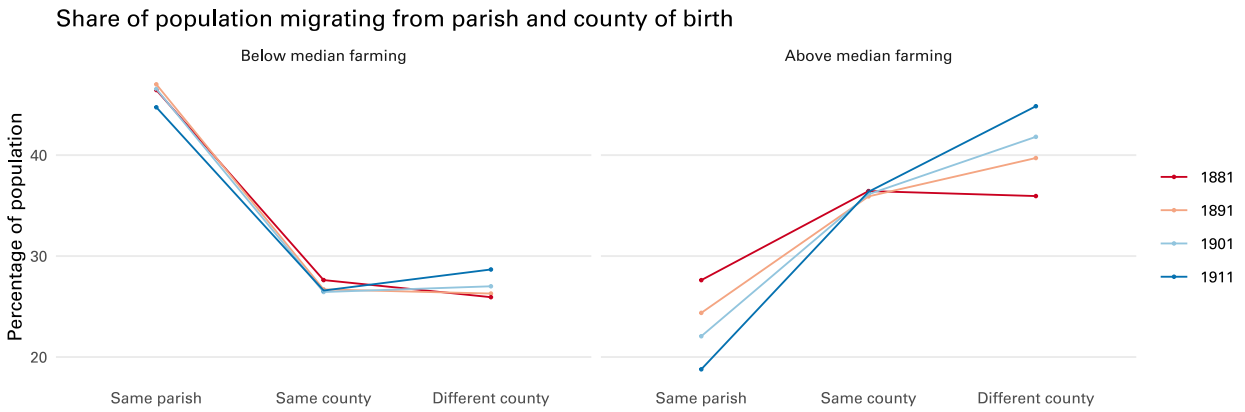


Figure 1: People migrated over longer distances, especially from rural areas

This figure plots the share of the population living in their parish of birth, in another parish of the same county, or outside the county of birth, subset by whether the parish of birth had an above- or below-median share of employment in farming in 1881.

psychological channel through which, all else equal, more culturally distant migration reduces identification with the working class.

3 THEORETICAL FRAMEWORK

We link migration-induced cultural distance to class identity using the framework developed by Shayo (2009; 2020). This framework draws on theoretical insights and experimental results from social psychology, which show that identification with social groups is malleable and governed by specific psychological mechanisms: people tend to identify with higher-status groups (Tajfel and Turner, 1986) and are more likely to see themselves as group members when they share similarities with that group (Turner et al., 1987). Shayo formalizes these insights as a choice problem, where the likelihood of identification increases with the group's status and decreases with the perceived distance between the individual and the group. Perceived distance is defined as the extent to which an individual's attributes diverge from those of the prototypical group member.

In our setting, individuals choose whether to identify with the working class (w) or with its complement, an upper- and middle-class identity (u), henceforth "upper." They reside in a given location and are characterized by their occupational class and their culture, which varies by region of origin.[6] Occupational class is objectively assigned—for instance, miners and shipbuilders would be classified as working class by census designations. By contrast, class identification is subjective: it is a choice that may or may not align with an individual's

[6]We use *culture* as a proxy term encompassing multiple features such as language, religion, and customs that result from socialization, vary across locations, and signal differences across groups of individuals.

actual occupation. The utility a member of culture k in occupational class c receives from identifying with class $j \in \{w, u\}$ is

$$v_{k,c}^{j} = \gamma\sigma^{j} - \delta d_{c}^{j} - \beta d_{k}^{j}. \tag{1}$$

Here σ^j is the average status of those with occupational class j. d_c^j and d_k^j capture an individual's perceived distance from class j based on different attributes.[7] d_c^j is the distance based on occupational class and captures whether a person views themselves as closer to the working or upper class given their occupation. d_k^j is the distance based on cultural origin.

To align the model with our empirical setup, consider an individual with either a working- or upper-class occupation who originates from one of K origins. In that case, $d_c^j = \mathbf{1}_{\{c \neq j\}}$: distance is zero if c is of the same occupational class as j, one if not. The individual's cultural distance to class j is then defined as the average cultural distance to members of that class in the individual's location. We can write that as a weighted average:

$$d_{k}^{j} = \sum_{r=1}^{K} s_{r}^{j} \left(1 - \varphi_{kr}^{j}\right)$$

where φ_{kr}^j denotes the cultural proximity within class j between an person from region k and one from region r. s_k^j is the share of members of class j originating from cultural region k, so that $\sum_{r=1}^{K} s_r^j = 1$ for each $j \in \{w, u\}$.

Inserting these expressions for distance based on occupation and on cultural origin into (1) and taking the difference between the payoffs from identifying with the working class ($v_{k,c}^w$) and upper class ($v_{k,c}^u$) gives:

$$v_{k,c}^{w} - v_{k,c}^{u} = \underbrace{\gamma\left(\sigma^{w} - \sigma^{u}\right)}_{\text{Relative status of working class}} + \underbrace{\delta\left(\mathbf{1}_{\{c \neq u\}} - \mathbf{1}_{\{c \neq w\}}\right)}_{\text{Occupational class}} \underbrace{-\beta \sum_{r=1}^{K} \left(s_{r}^{w}(1 - \varphi_{kr}^{w}) - s_{r}^{u}(1 - \varphi_{kr}^{u})\right)}_{\text{Relative class distance}} \tag{2}$$

Individuals will choose to identify with the working class when the status of the working class is higher relative to the upper class ($\sigma^w - \sigma^u$ is large), when they themselves belong to a working class occupation ($c = w$) and when the share of people from culturally similar origins

[7] The cultural choice problem in Shayo (2020) is:

$$\max_{J \in G_i} \{\pi_i - \beta_i d_{iJ} + \gamma_i S_J\}$$

where π_i is material payoff, d_{iJ} distance of i to group J, and S_J is the status of group J. Our formulation differs from Shayo's in dropping the π_i term, which is irrelevant for cultural choice, and in separating distance into horizontal and vertical components.

is higher among the working class relative to the upper class ($\sum_{r=1}^{K}(s_r^w(1-\varphi_{kr}^w) - s_r^u(1-\varphi_{kr}^u))$ is negative).

Equation 2 can be directly expressed as a regression equation for empirical estimation. Conditioning on the relative status of classes and an individual's occupation, the framework predicts that β should be negative: people are less likely to identify with the working class when its average cultural distance, relative to that of the upper class, is higher.

What does identification with the working class imply? In Shayo (2009), identifying with a group means holding different preferences over outcomes. As a result, group identity leads individuals to deviate from pure self-interest in two ways: first, by internalizing the group's welfare, people may pursue goals that benefit the group even at personal cost; second, they may adjust their behavior to conform to group norms. In our context, identification with the working class may manifest both in actions that promote working-class interests—such as voting for a party whose policies favor the working class—and in individual choices that signal conformity with working-class tastes, such as consumption patterns or leisure activities. Our empirical measures—naming choices, union membership and voting—capture both manifestations of class identification. The next two sections outline how we measure identity choice and relative cultural distance, and how we estimate the effect of the latter on the former.

4 Measuring class identity and cultural proximity

4.1 *Data*

We link cultural distance from different classes with individual identity choices by examining the names that parents gave their children as the composition of their place of residence changed over time. We rely on individual census microdata from the full count of the Census of England and Wales, 1881–1911 (Schürer and Higgs, 2024), to measure class identity through names and to construct demographic composition at the parish level. Below, we explain the construction of each of these measures in more detail.

4.2 *Measuring class identity*

We measure class identity at the individual level using first names. The basic idea is that some first names are more common in working-class occupations and others in middle- and upper-class ones. While many factors shape naming decisions, identification with a specific class should also be reflected in those choices.

Our measure of the class valence of name i is

$$\textit{Working Class Score}_i = \frac{P(\text{name} = i|\text{working class})}{P(\text{name} = i|\text{working class}) + P(\text{name} = i|\text{upper class})} \times 100.$$

This score corresponds to the probability a Bayesian with a flat prior would assign to an individual with name i being working class. Such scores have been developed to measure the distinctiveness of Black names (Fryer and Levitt, 2004) and proxy for national, ethnic and local identities (Abramitzky, Boustan and Eriksson, 2014; Fouka, 2019; Bazzi et al., 2019; Fouka, 2020; Abramitzky, Boustan and Eriksson, 2020), but have not, to the best of our knowledge, been used to measure class identity.

We construct these scores using the first names of people of working age (16–60) recorded in the 1881 census. We allocate men to classes based on their recorded occupation. For women, we use the class of their husband. The *Seventy-fourth Annual Report of the Registrar-General* (1913) provides the earliest allocation of occupations to social classes. The Registrar-General allocated occupations to eight groups, with 1 corresponding to "the upper- and middle-class," 3–8 to "the 'working-class'—those employed in manual labour, skilled or unskilled" and 2 "intermediate between the middle and working class because it consists of occupations, such as the shopkeeping trades, including many members of both classes" (p. xli). We classify those occupations in class 1 as upper class and those in classes 3–8 as working class. A major advantage of using this classification is that it provides information on the class valence of different occupations as they appeared at the time. We standardize names using Metaphone codes and drop names that appear fewer than 100 times in our sample.

Table A.1 lists the highest and lowest-scoring names. These names have clear class valences. Distinctively working class names are often informal contractions, such as Pat, and tend to have biblical origins, for instance Esau, Noah, Job, and Enock. Working class names like Patrick and Bridget trace out Irish immigration to England and Wales in this period. Low-scoring names like Cecil, Gerald, and Algernon feature as the names of characters in plays by Oscar Wilde. Upper class names tend to have classical origins; the lowest scoring, Cyril, originates in the Greek word for "lord." Tables A.2 and A.3 report the occupations with the highest and lowest average scores in the 1911 census. Various forms of textile, shipyard, and mining laborers dominate the top of the list. Occupations with the lowest scores are students—in a period when a tiny minority received higher education—military officers, bankers, and lawyers.

Empirically, the names people gave their children were strongly associated with their own occupational class. Figure 2 plots the negative relationship between an occupation's status, estimated by the HISCAM project, and the average working class score of children of men

in that occupation, in the 1911 census.[8] Hotbeds of working class identity like mining are above the regression line, while clerical occupations, which were not particularly high status but which did not involve manual labor and were considered lower middle class, are below the regression line. Table A.4 shows that this relationship holds even when comparing the children of fathers with the same first name (who have the same working class score).

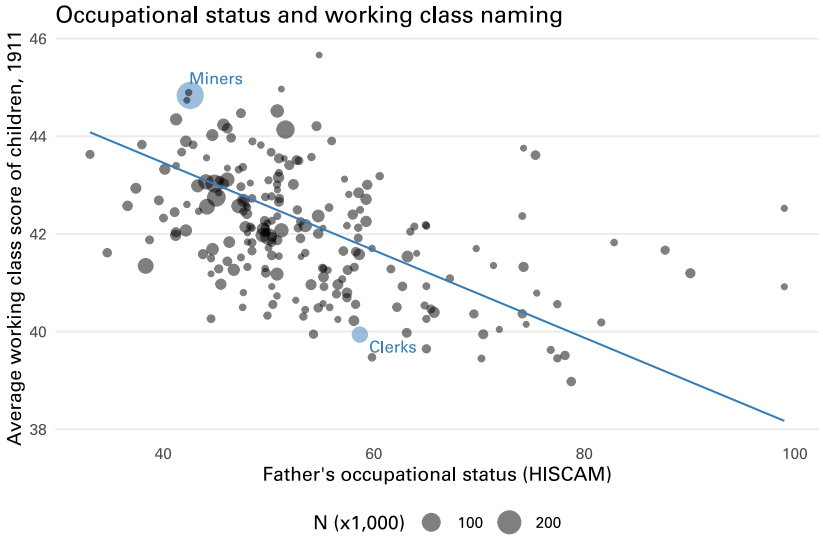


Figure 2: Children's name scores correlate with occupational status

This figure plots the average working class score for children of men with in a given HISCO occupation in the 1911 census, against the HISCAM occupational status score for that occupation. The size of dots corresponds to the number employed in each occupation. The blue line shows the OLS fit.

To more systematically validate this measure, we look at how it correlates with other class markers among a single occupation: Members of Parliament. Rush (2001) collected detailed information on the backgrounds and social lives of members of parliament, from contemporary publications such as parliamentary guides. This data has subsequently been distributed by History of Parliament Trust and Rush (2020). This information includes factors like aristocratic affiliation, education, and social club memberships. Table 1 shows the relationship between these class markers and working class scores, among MPs born between 1850 and 1920. Those with aristocratic connections or knighthoods, alumni of public schools (elite private schools) or Oxford and Cambridge ("Oxbridge"), and members of exclusive social clubs, had lower working class scores on average. Labour MPs—in a period when an aim of the Labour Party was to elect working men to parliament—and MPs affiliated with trade unions, had higher working class scores on average. These correlations increase our

[8]HISCAM is a measure of occupational status constructed from marriage data from the same era under the assumption that people tend to marry people with similar social status (Lambert et al., 2013).

	Working Class Score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Aristocratic connections	-2.002^{*}						-0.594
	(0.783)						(0.855)
Knighthood	-2.094^{**}						$-1.314^{\dagger}$
	(0.641)						(0.672)
Labour		3.417^{**}					2.230^{*}
		(0.646)					(1.092)
Liberal		-0.282					-0.932
		(0.706)					(0.732)
Public school			-2.820^{**}				-1.887^{*}
			(0.555)				(0.743)
Oxbridge				-1.958^{**}			0.041
				(0.627)			(0.775)
Other university				$-1.328^{\dagger}$			-0.607
				(0.734)			(0.766)
Social club member					-2.209^{**}		-0.649
					(0.549)		(0.662)
Union affiliate						2.614^{**}	-1.136
						(0.656)	(1.034)
N	2951	2951	2951	2951	2951	2951	2951
R^2	0.016	0.020	0.019	0.014	0.016	0.015	0.026

This table presents predictors of MPs' class scores. The sample is restricted to MPs born between 1850 and 1920, who were members of one of the three main parties. For the party variables, the reference category is Conservative. All models include a third-degree polynomial in birth year and an intercept. Robust standard errors in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table 1: Predictors of class scores among MPs

confidence that the content of first names, as measured by working class scores, captures a meaningful expression of class identity.

4.3 *Measuring cultural distance from the working class*

Our main hypothesis is that class identity, as expressed through the first names chosen for one's children, should depend on how culturally similar people in that occupational class are in the father's location of residence, relative to members of the other occupational class (Equation 2). Using birth parishes as our measure of regional origin, we calculate the direct empirical analogue of the *Relative class distance* in Equation 2, for each origin parish-residence

parish-year combination:

$$\text{Relative class distance}_{ipt} = \underbrace{\sum_{l} s^{w}_{lpt}(1-\varphi^{w}_{il})}_{\text{Distance to working class}} - \underbrace{\sum_{l} s^{u}_{lpt}(1-\varphi^{u}_{il})}_{\text{Distance to upper class}} .$$

Here s^{w}_{lpt} and s^{u}_{lpt} is the share of adult men with working or middle and upper class occupations in parish p at time t who were born in parish l. φ^{w}_{il} and φ^{u}_{il} are measures of cultural proximity between the father's birthplace i and the resident's birthplace l for the working and upper class.

Throughout, we use the consistent parish units created by Schürer and Higgs (2024) as the geographical unit. These consistent parishes were created by combining together parishes that were subsequently merged and so correspond to realistic places. We geocode consistent parishes by merging the 1911 census to a GIS of parishes in that year from the Great Britain Historical GIS and aggregating 1911 parishes into consistent parishes.

The challenge is to quantify cultural proximity (φ^{w}_{il} and φ^{u}_{il}) between individuals from two birth locations, i and l. Cultural variation in Britain stemmed from distinct geographies, histories, and patterns of communication over time (Homans, 1969), and was expressed in differences of speech, customs, and religion. These differences were not only location-specific but also class-specific. For instance, in parts of Wales, the upper class tended to be Anglican and English-speaking, and so culturally closer to people in England, while the working class tended to be Welsh-speaking and Nonconformist.

To capture cultural distance across locations and classes, we use historical data on intermarriage. Patterns of intermarriage reflect multiple forms of compatibility between people in different locations, including shared religion and similar customs. For any two parishes i and l, we calculate

$$\varphi^{j}_{il} = \frac{m^{j}_{il}}{\sqrt{m^{j}_{i} m^{j}_{l}}},$$

where m^{j}_{il} are the number of married couples of class $j \in \{u, w\}$ in which one spouse is born in i and the other in l, and m^{j}_{i} and m^{j}_{l} are the total number of married couples of class j with at least one born in i and l respectively. This variable is the geometric mean of the probability of someone born in i marrying someone born in l and the probability of someone born in l marrying someone born in i. We calculate this variable using information from the 1881 census, the earliest in our data, inferring the class of couples from the husband's occupation. Allowing this measure to vary by class means it can pick up cases in which the proximities of the upper and working classes between two locations differ. Figure A.2 shows marriage

connections in 1881 for the upper and working classes for the parishes of Manchester and Plymouth. Connections are higher to nearby locations, but this pattern is stronger for the working class. Members of the upper class are more likely to marry across larger distances and across religious divides.

Table A.5 shows the correlates of φ_{il}^{j}. In general, parishes that are geographically closer have higher rates of intermarriage. However, conditioning on geographic distance, other forms of cultural distance, such as belonging to different counties, differences in the popularity of religious denominations as recorded in the 1851 religious census, and differences in the frequencies of surnames in the 1861 census are negatively correlated with intermarriage.[9] While we obtain the same pattern of coefficients for both upper and working class intermarriage, the coefficients on all forms of distance are substantially larger for the working class. This pattern makes sense given greater cultural heterogeneity within the working class. For example, consistent with greater proximity to the English among the Welsh upper class, as compared to the Welsh working class, we estimate a smaller effect of two locations being in the same country on upper class marriage connections. We rely on cultural distance based on intermarriage rates in our main analyses, but we demonstrate robustness to alternative proxies for cultural distance.

5 INDIVIDUAL-LEVEL IDENTITY SHIFTS

5.1 *Empirical strategy*

Our sample consists of all children aged 20 or younger living in the same household as their father in the 1881, 1891, 1901, and 1911 censuses. We assign children working class scores using their first names, as explained in section 4.2. The central idea is that the working class score of a child born in a given parish and year reflects the father's class identity in that place and time, which would have been shaped by the parish's cultural composition.[10] Examining how working class scores of children change over time *within* the same household should isolate the effects of changing parish cultural composition, holding constant other determinants of class identity, such as the father's social background.

Our analysis leverages two sources of identifying variation: relative class distance can change across children in the same household over time either because the household itself moves to a different parish, or because the composition of the parish changes due to migration.

[9] We use the 1861 census for surnames because it is the earliest census for which we have access to usable data on individuals' names.

[10] Clearly, children's names reflect the class identity of both parents, but men were more likely to participate in the labor force during the period of study, and we therefore rely on the father's occupation to condition on (objective) class.

The top panel of Figure 3 illustrates the first source of variation. It shows two households in the 1901 census. George Shean and William Gibbon, both born in Bath, each had two children, one born in 1898 and one in 1900. Their first children—Olive Shean and Doris Gibbon—were both born in Bath, Somerset, and receive similar working class name scores (41 and 39). Between 1898 and 1900, George Shean moved to London, and his second child, Edith, was born in Hammersmith. Because individuals born in Somerset are more culturally distant from the working class in London than in Bath, this move increased George Shean's relative class distance from -0.11 to 0. This shift is reflected in naming choices: Edith's name receives a working class score of 21, a decrease of 20 relative to her older sibling. In contrast, William Gibbon did not move. His second child, Olive, was also born in Bath, implying no change in relative class distance. Consistent with this, his children's name scores remain stable, changing only from 39 to 41.

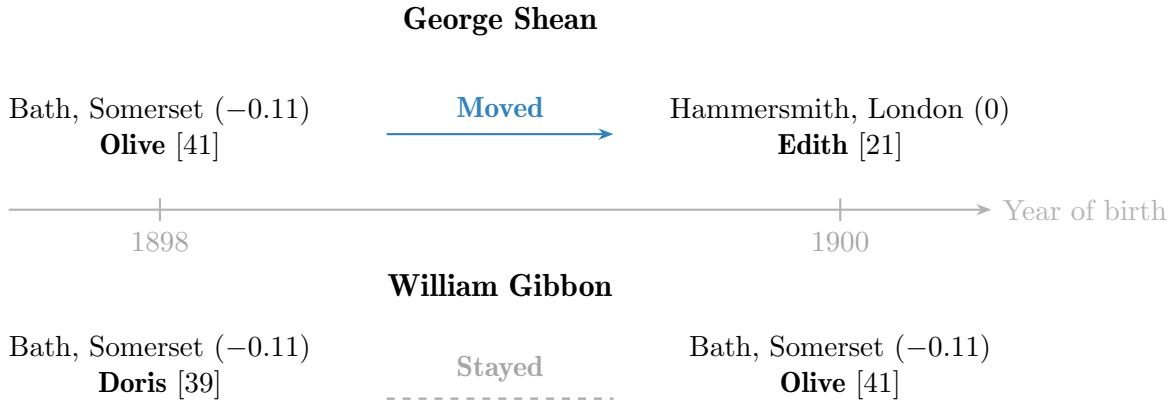
The bottom panel of Figure 3 illustrates variation from individuals who do not migrate but whose relative class distance changes due to migration into their place of residence. Joseph Froman and James Atwood were also recorded in the 1901 census and had 2 children. Both lived in their parishes of birth; Rowley Regis in the case of Froman, and Kingswinford in the case of Atwood. Both parishes are in Staffordshire, near the town of Dudley. Both men had children born in 1894 and 1900, and gave their first children names with similar class valences. Froman named his first child Ellen, which receives a score of 52, while Atwood named his first child Annie, which receives a score of 53. Between 1891 and 1901, Rowley Regis drew in a range of working class migrants. The number of working class male residents born in other parishes increased from 3,087 to 4,388. This population change included an increase in the number of working class men born in London from 60 to 519. As a result, relative class distance for a resident of Rowley Regis born in the parish increased from -0.12 to -0.071. In contrast, the number of working class migrants in Kingswinford was stable, increasing from 3,321 to 3,583, corresponding to a change in relative class distance from -0.095 to -0.080 for residents born in the parish. These changes in relative class distance are reflected in the names Froman and Atwood gave their children. Froman's second child, Beatrice, receives a working class score of 18, a decrease of 34 from the first child, while Atwood's second child, Martha, receives a score of 55, an increase of 2 from the previous child.

Our baseline estimating equation is as follows:

$$\textit{Working class score}_{ipt} = \beta\, \textit{Relative class distance}_{ipt} + \gamma_i + \delta_t + \varepsilon_{ipt}. \tag{3}$$

where i indexes fathers, p parishes of birth, and t years of birth, so $\textit{Working class score}_{ipt}$ is the working class score of a child of i born in p at time t. We assign $\textit{Relative class distance}_{ipt}$

(a) Household moves to a different parish



(b) Parish class composition changes due to migration

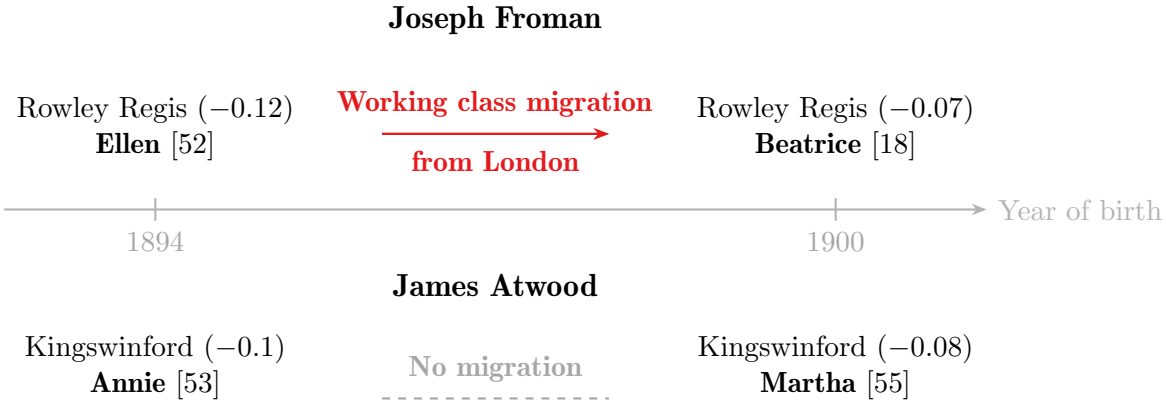


Figure 3: Examples of identifying variation

Notes: George Shean, William Gibbon, Joseph Froman, and James Atwood are observed in the 1901 census. Numbers in parentheses denote the relative class distance of the children's parish of birth; numbers in brackets denote the working class score of children's names.

to a child's parish of birth p based on the parish's demographic composition in the census year nearest to the child's year of birth t. γ_i is a fixed effect for the father and δ_t is a year of birth fixed effect, which absorbs any cohort trends in naming patterns that affected children in all parishes. To add precision—although we omit these variables from the estimating equation for simplicity of presentation—we control throughout for the child's sex and birth order. We also estimate more parsimonious specifications with additional fixed effects, which we discuss as they become relevant.

Our design is a generalized difference-in-differences with continuous treatment. It rests on two identification assumptions. The first is parallel trends at the level of fathers: fathers who experience changes in relative class distance would have otherwise followed similar trajectories in terms of class identity. We provide evidence in support of this assumption by showing that households experiencing changes in their parish's relative class distance were not already on a trajectory of more working-class naming for children born prior to the change. The second assumption is that no other shocks affecting class identity or naming occur contemporaneously with changes in relative class distance. To support this, we show that our results are robust to specifications that include parish by year of birth fixed effects. These specifications hold constant all factors that vary over time at the parish level—including determinants of class identity such as migration, economic conditions, and industry composition—and identify the effect of relative class distance only from variation among fathers of different cultural origins within the same parish.

Equation 3 with parish by year of birth fixed effects is the empirical counterpart of equation 2, which specifies the relative payoffs to working vs upper class identification. The other theoretical determinants of this identification are absorbed by the fixed effects: father fixed effects account for occupation, while parish fixed effects capture the (possibly time-variant) relative status of classes within each location. Conditional on these factors, the relative appeal of working class identity as reflected in naming choices is a function of relative cultural distance from the working class.

5.2 *Results*

Our main estimates are presented in Table 2. Column (1) shows that greater relative class distance is associated with less working class naming. This effect reflects both distance from the working class and proximity to the upper class: as shown in column (2), distance from the working class decreases working class naming, while distance from the upper class is associated with higher working class scores. Columns (3) and (4) add parish fixed effects, which account for local naming preferences and conventions affecting all children born in a given parish. Columns (5) and (6) include parish-by-year-of-birth fixed effects. Although

this specification substantially limits identifying variation—now coming only from fathers of different origins who experience different changes in relative class distance within the same parish—the effect of relative class distance remains large and significant.

Columns (7) and (8) estimate an alternative specification that includes fixed effects for the child's birth parish interacted with fixed effects for the father's birth parish. This specification also exploits variation across fathers of different origins experiencing different changes in class distance, but in this case over time: it accounts for all factors that affect naming choices specific to each origin–destination pair and isolates the impact of changing relative class distance within a given father-origin–child-birthplace cell. This specification accounts, for instance, for naming choices common to all individuals born in Manchester and living in London; identification comes only from changes in the composition of their residence parish over time. The effect of relative class distance remains significant and is primarily driven by distance from the working class.

Model (1) implies that moving from the 5th to the 95th percentile of relative class distance corresponds to a 0.12-unit decrease in working class scores, while the equivalent decrease for model (7) is 0.21. These predicted effects are sizable, corresponding to 6% and 11%, respectively, of the difference in working class scores between Oxbridge-educated MPs and non-graduate MPs in Table 1.

To assess how each of our sources of identifying variation—households moving across parishes and parishes changing composition due to migration—affects class identity we estimate equation (3) in first differences:

$$\Delta \textit{Working class score}_{ipt,p't'} = \beta \Delta \textit{Relative class distance}_{ipt,p't'} + \delta_t - \delta_{t'} + \Delta \varepsilon_{ipt,p't'}. \quad (4)$$

Here $\Delta \textit{Working class score}_{ipt,p't'}$ is the change in working class scores between i's children born in p at time t and the preceding child, born in p' at t'. This specification allows us to isolate movers and separately estimate effects for them and for families that remain in the same parish across consecutive births but experience changes in parish class composition. It also enables us to examine within-household trends in naming across births, providing evidence in support of the parallel-trends assumption underlying our design.

Figure 4 illustrates this evidence. It reports estimates from separate regressions based on Equation 4, in which within-household changes in working-class scores are regressed on future, current, and past changes in relative class distance, including the full set of parish fixed effects from Table 2. The x-axis indexes time across two consecutive births. Changes in relative class distance between births at $t = 0$ and $t = -1$ are uncorrelated with changes in the working class score of children born at $t = -1$ and $t = -2$. This indicates that households

	Working class score							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative class distance (working − upper)	−0.801**		−0.819**		−0.520*		−1.444*	
	(0.212)		(0.224)		(0.213)		(0.563)	
Distance to working class		−0.921**		−0.852**		−0.675**		−1.805**
		(0.237)		(0.250)		(0.257)		(0.599)
Distance to upper class		1.162**		0.918*		0.934*		0.661
		(0.385)		(0.400)		(0.423)		(0.660)
Fixed efffects: Household	x	x	x	x	x	x	x	x
- Year	x	x	x	x			x	x
- Parish			x	x				
- Parish × year					x	x		
- Parish × father's parish of birth							x	x
Controls: Sex, Birth order	x	x	x	x	x	x	x	x
R^2	0.469	0.469	0.470	0.470	0.482	0.482	0.504	0.504

This table presents evidence of the relationship between distance from the working or upper class and working class naming. Data is at the level of births, differentiated by father, year, and parish of birth. The dependent variable is the Working Class Score for the child. All models control for sex and birth order. Standard errors clustered by household and parish in parentheses. The number of observations in all models is 25,773,095. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table 2: Within-father relationship between relative class distance and working class naming

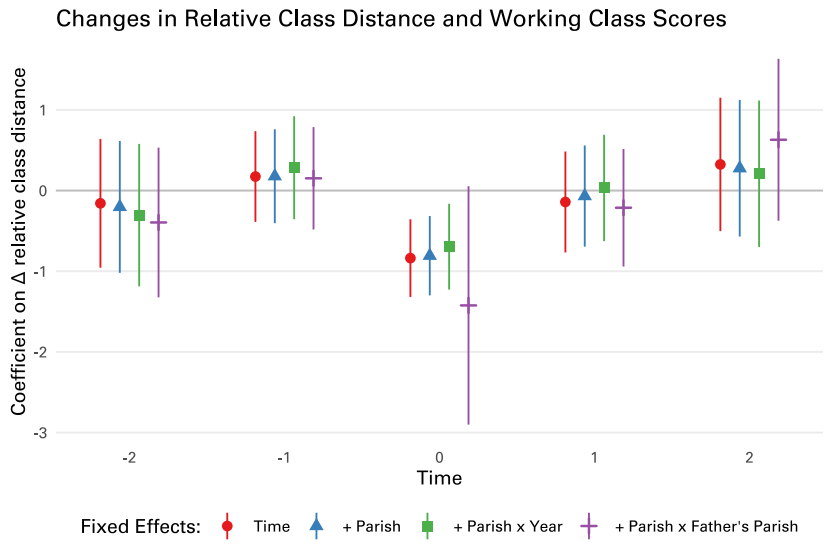


Figure 4: Dynamic relationship between changes in relative class distance and naming

Each point is the coefficient from a separate regression of the within-father first difference in working class scores against the within-father lead or lagged change in relative class distance. Red dots are from models with fixed effects for present and lagged year of birth, blue triangles also include fixed effects for present and lagged parishes of birth, green squares for present and lagged parish-by-year of birth, and purple crosses present and lagged parishes of birth interacted with the father's parish of birth. Error bars show 95% confidence intervals, calculated from standard errors clustered by father and present and lagged parish. Each coefficient is from a separate regression.

that moved to locations with greater relative class distance were not following a differential trend in working class naming prior to the move, nor were those in locations in which working class distance subsequently increased due to migration. We also find that the effect of a change in relative class distance does not persist to subsequent births. This pattern implies a level shift: higher distance from the working class lowers the working class score of the next child, which then remains at this lower level for later births rather than continuing to fall.

Figure 5 displays our main results and checks for the parallel trends assumption separately for the two groups of households that drive our identifying variation: those that changed parishes between the present and previous child ("movers"), and those that did not ("stayers"). The first row plots estimates from regressing the change in working class scores between children in a household against the change in relative class distance as in equation (4). The estimated effects for movers and stayers are similar. The coefficients for movers are more precisely estimated reflecting more within-household variation in relative class distance among the households that move. The second row regresses the change in working class scores between the previous two children against the change in relative class distance. We find no evidence of pretrends for either movers or stayers.

Finally, we regress pre-change household characteristics on subsequent changes in relative

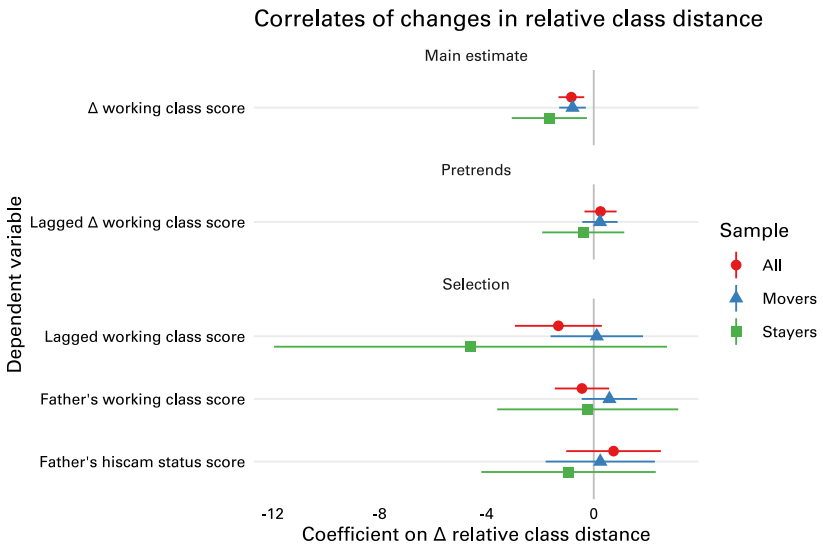


Figure 5: Within-household first difference estimates: pretrends and selection

This figure reports coefficients from within-household first-difference regressions of changes in outcomes against changes in relative class distance between births. The first row shows the main estimate in first-difference form, regressing the change in children's working class scores against changes in relative class distance, with fixed effects for the current and lagged year of birth, controls for gender, and 95% confidence intervals constructed from standard errors clustered by household, current parish and lagged parish. The second row shows the relationship between changes in relative class distance and changes in working class scores between the previous two children. The next rows show the relationship between changes in relative class distance and the working class score of the previous child, , the father's own working class score, and a measure of the father's occupational status. Red dots use the entire sample, blue triangles subset to households who change parishes between the current and previous birth, green squares subset to households who do not change parishes.

class distance. Although our design requires only parallel trends and not identical pre-change levels of covariates between households that did and did not experience a change in relative class distance, the plausibility of this assumption is strengthened if those moving to (or remaining in) places where they became culturally more distant from the working class were otherwise similar. We find no evidence that prior patterns of working-class naming, the father's occupational status, or his working class name score are correlated with subsequent decreases in relative class distance. This absence of selection into less working-class locations reinforces the credibility of our identification strategy.

5.3 *Robustness and alternative explanations*

Appendix B presents additional evidence of the robustness of the results and investigates alternative mechanisms. We show the results are unchanged using alternative measures of cultural proximity (Table B.1), and measures of class identity specific to the period (B.2) or region (B.3). Our results are not driven by specifically Irish names (B.4), nor by

the particularities of the sample of households with multiple children that provide within-household variation. Re-estimating the models with less-granular fixed effects gives larger coefficients (B.5), consistent with a restrictive within-household specification exacerbating measurement error and biasing coefficients towards zero.

We find evidence inconsistent with alternative mechanisms. A prominent one is aspirations for social mobility, which may lead parents to choose more middle class names for their children if they observe more culturally similar individuals to them in that group, and therefore infer that social mobility is attainable for their offspring. If this mechanism were a crucial driver of our results we would expect stronger naming responses for households already closer to the middle class. Yet we find no heterogeneity across households of different occupational statuses (Table B.6 and Figure B.1). We also consider the possibility that relative class distance may capture distance in occupation—people from culturally distant backgrounds might sort into different occupations, with the resulting distance dampening working class identity. While we find that distance in occupation matters for working class identification—a mechanism conceptually similar to ours—relative class distance remains predictive (Table B.7). Finally, we address the concern that lower working class identification may be confounded with higher national identification, as in Shayo (2009). We construct a measure of local identity analogous to our working class scores, capturing the relative frequency of names in each county compared to the overall population. While higher distance from the working class reduces local identification (Table B.8), consistent with an increase in national identity, this channel does not explain our main results. Working class scores are uncorrelated with local identity (Table B.8) and flexibly controlling for local identity leaves our main estimates unchanged (Table B.9).

6 INDIVIDUAL-LEVEL POLITICAL BEHAVIOR

The previous section established that relative class distance lowered identification with the working class as reflected in name choices. Did these individual identity changes matter for political behavior? To address this question, we analyze individual-level data on trade-union membership: a central individual expression of working class political identity at the turn of the twentieth century.

6.1 *Data*

We use individual-level data on trade union membership during the period 1871–1921. Union registration and admissions books are held at the Modern Records Centre at the University of Warwick, and the subset we rely on has been digitized by the genealogical company

Findmypast. This digitized collection covers fourteen large unions.[11] Although our data do not include the universe of unions, our empirical strategy allows us to study the decision to join a union among individuals in occupations covered by the unions for which we have data, yielding internally valid inferences for this subsample.

Our data include each member's first and last name, the union they joined, the location of their union branch, and their date of registration or admission. After removing duplicate records, we are left with 1,926,306 individuals across 5,630 union branches. Figure A.3 shows the geographic distribution of these branches, by union and membership size.

We link these union records to census microdata. Our goal is to identify how relative class distance across parishes and occupations affected the probability of joining a union. A key challenge is that information is not consistently available across observations; for instance, some records list full names while others provide only initials. To address this, we use the probabilistic merging algorithm developed by Enamorado, Fifield and Imai (2019). We restrict matches to census observations within ten years of the union record, employed in occupations relevant to the corresponding union, and residing in the same county. The key advantage of the probabilistic algorithm in our context is that it provides a principled way to account for missingness in the data. Our results are robust to using a deterministic algorithm, similar to Abramitzky, Mill and Perez (2020), which requires exact matches on geography (parish or county), sector of employment, and various formulations of names (full names, initials, and metaphone codes).

Our premise is that union membership and naming patterns both constitute expressions of the same latent class identity. In support of this interpretation, we find that, among people employed in the same occupation, union members give their children more working class names. Table A.6 shows that, with occupation-by-year fixed effects, the gap in children's name scores between union and non-union members is about half as large as the gap between Oxbridge and non-graduate MPs.

6.2 *Empirical analysis*

Linking union members to census records allows us to examine whether relative class distance, measured among people in the same occupation and location, predicts the likelihood of joining a union. Throughout, we restrict the sample to individuals employed in occupations relevant

[11]The unions included are Boilermakers & Iron Shipbuilders, Bricklayers, Carpenters & Joiners, Compositors, Correctors of the Press, General Workers, Insurance Workers, Lithographers, Local Government Officers, Paper Makers, Printers, Railway Workers, Teachers, and Typographers.

to the unions in our data. Our preferred specification has the following form:

$$\textit{Union member}_{it} = \beta \textit{Relative class distance}_{ipt} + \gamma_{pot} + \delta_{bot} + \varepsilon_{it} \tag{5}$$

The dependent variable, $\textit{Union member}_{it}$, equals 1 if individual i appears in the union membership records at time t, and 0 otherwise. The key independent variable, $\textit{Relative class distance}_{ipt}$, is the relative class distance for i and are resident in parish p at time t. γ_{pot} is a fixed effect for the parish of residence-occupation-year combination specific to i, which absorbs all factors that affect unionization rates for a given occupation in a given place and year. δ_{bot} is a fixed effect for the parish of birth-occupation-year combination specific to i, which absorbs all time-varying factors associated with people born in a given place that might independently affect their propensity to join a union in that occupation. The coefficient β is therefore identified by comparing individuals in the same occupation and born in the same parish who live in different locations—and thus face different relative class distance—while holding constant baseline unionization rates in each residence-parish–occupation cell and the time-varying baseline propensity of each birth-parish–occupation group to unionize.

Table 3 presents our results. We progressively build toward our preferred and most restrictive specification in equation 5. Model (1) includes only occupation-by-year fixed effects; the estimate therefore reflects the within-occupation correlation between relative class distance and union membership. The negative coefficient indicates that moving from the 5th to the 95th percentile of relative class distance corresponds to roughly a one–percentage-point decrease in the probability of union membership—about 8% of the baseline unionization rate in our data. Model (2) adds occupation-by-parish-by-year fixed effects, holding constant location-specific factors that might affect unionization within each occupation. Model (3) implements our preferred specification by further adding birth-parish–by-occupation–by-year fixed effects, and Model (4) adds controls for age and gender.

Models (5) and (6) restrict the sample to those in working class and upper or middle class occupations, respectively. The estimates are driven by individuals in working class occupations. This pattern is reassuring: relative class distance should shift class identity away from the working class, and mobilizing on behalf of one's occupation is a meaningful expression of working class identity only for those actually employed in working class jobs.

Our analysis leverages variation in relative class distance across residence locations among people from the same origin and in the same occupation. But working class distance may also vary across occupations. To account for this, we construct within-occupation measures of cultural distance. Table A.7 shows that cultural distance to others in the same occupation is negatively correlated with union membership, but this relationship disappears when

	Union member					
	(1)	(2)	(3)	(4)	(5)	(6)
Relative class distance	-0.079^{**}	-0.046^{**}	-0.060^{**}	-0.030^{**}	-0.034^{**}	0.004
	(0.030)	(0.010)	(0.010)	(0.009)	(0.011)	(0.003)
Fixed efffects: Occupation × year	x					
- Occupation × parish × year		x	x	x	x	x
- Occupation × birth parish × year			x	x	x	x
Controls: Gender, age				x	x	x
Sample:	All	All	All	All	Working	Upper
N	6052902	5930345	5707463	5692870	4951251	688950
R^2	0.281	0.409	0.465	0.467	0.454	0.278

This table presents evidence of the relationship between relative distance from the working class and union membership. Each observation is an individual worker in a given census year, for the period 1881–1911. The sample is restricted to those employed in occupations related to the Carpenters and Joiners, Railway Workers, Teachers, Boilermakers and Iron Shipbuilders, Typographers, Lithographers, Bricklayers, Papermakers, Printers, Local Government Officers, Insurance Workers, and Compositors unions. The dependent variable takes a value of 1 if the worker is a member of one of these unions, 0 otherwise. Models (4)–(6) control for gender and a third-degree polynomial in age. In (5) the sample is further restricted to those in occupations classified as working class by the Registrar-General, in (6) to those classified as upper or middle class. OLS estimates. Standard errors clustered by parish and birth parish in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table 3: Relationship between class distance and union membership

controlling for distance to the working class residents of the same parish.

Our findings are not sensitive to the method used to link union and census records. Table A.8 reproduces our main results using a deterministic linking algorithm. The estimates are very similar under this more conservative approach, indicating that false positives do not drive our findings. We also rule out the possibility that our results are an artifact of linkage biases. The concern is that relative class distance might predict the probability of a successful merge rather than union membership per se. While we cannot test this directly in the census data, we can examine what predicts whether union records are successfully merged to the census. Table A.9 shows that the probability of a successful merge is uncorrelated with an individual's working class score, making it unlikely that class identity or its expressions—such as naming patterns documented earlier—affect linkage probabilities. Instead, merge success correlates with mechanical factors such as first-name length (records are harder to merge when the first name is only an initial) and the number of individuals sharing the same first name. Controlling for these predictors of linkage success leaves our results unchanged (Table A.10).

7 AGGREGATE POLITICAL EFFECTS

In this section, we link individual-level choices to aggregate political outcomes. We show that local-level distance from the working class affected not only individual identity and the decision to join unions, but also the electoral success of the Labour party and the rhetoric of parliamentary candidates. We then examine the spatial heterogeneity of changes in relative class distance to better understand where and why support for Labour was favored by local demographic change. We show that urban areas experienced a decline in relative class distance between 1891 and 1911, as local population growth outpaced in-migration. This, in turn, facilitated increased support for Labour in cities, contributing to the emergence of an urban–rural political divide.

7.1 *Outcomes*

ELECTORAL SUCCESS OF LABOUR Since its inception as the LRC in 1900, the Labour Party set out to represent working class interests. Its rise marked the culmination of a restructuring of British politics from cleavages around religion (Wald, 1983) and tariff policy to one centered on class. If, holding fixed the occupational mix in an area, a shift in relative distance away from the working class made voters identify less with the working class, we would expect them to place less weight on working class interests and be less willing to vote for the Labour Party. This expectation aligns with Shayo (2009), which views group identity as leading individuals to adopt behaviors that advance group interests.

We focus our attention on three election years: 1892, 1900 and 1910. We do so because our independent variable, relative class distance, as well as other time-varying characteristics of constituencies, are derived from the census and vary at the decade level. Accordingly, we assign each election to its nearest census year in order to link constituencies to information on relative class distance and other controls.

During the study period, Labour's parliamentary representation rose from 2 seats in 1900 to 40 in 1910. A key factor in this breakthrough was a 1903 pact with the Liberal Party, under which the Liberals agreed not to contest certain constituencies (Adelman, 1996, p. 36). These constituencies were selected to maximize Labour's chances of winning. Moreover, because Labour was financially constrained, it concentrated on contesting seats in constituencies with a large working-class presence (Jusko, 2017).

ELECTION ADDRESSES Election addresses provide further information on the issues prioritized by voters. Parliamentary candidates could mail, free of charge, a leaflet describing their views on policy. These election addresses have been digitized by Daniel Ziblatt and Laura

Bronner. Our assumption is that candidates would tailor their addresses to the demands of their constituents. Middle and upper class voters had different preferences to working class voters, and so a candidate running in an upper-class area would emphasize a different set of issues, and use a different vocabulary, to a candidate running in a working class area. In our context, if a change in working class distance in a given area led voters to identify less with the working class, we would expect candidates to correspondingly emphasize working class issues less.

This assumption is justified. Hobsbawm (1984*b*, p. 208) notes that politicians were responsive to workers' class consciousness and "had to pay [their] respects to the supremacy of class when appealing to workers on the grounds of party." It is therefore reasonable to assume that they adjusted their appeals to different constituencies depending on the strength of working class identity. Bronner (2019) further emphasizes that candidates enjoyed significant autonomy in determining the content of their manifestos without being strictly bound to party lines, and shows empirically that manifesto content responded to constituency-level changes, such as the degree of enfranchisement.

We use supervised machine learning to create a measure of the extent to which candidates spoke to working or middle and upper-class concerns. Our basic approach is to "predict" the share of working class relative to upper class residents in a constituency using the frequency of terms used by candidates campaigning there. Our resulting measure is the predicted share of working class relative to upper class residents at the candidate-constituency-level. To construct this measure, we tabulate the frequency of all stemmed 1–3 word phrases in each election address, after stripping out stopwords, placenames from a gazetteer, and terms appearing fewer than 10 times. We also exclude terms that are only used in a given parliamentary constituency. Doing so helps to remove terms like the street addresses of candidates that are not theoretically relevant. Given these frequencies, we use Multinomial Inverse Regression (Taddy, 2013) to construct the measure. This routine involves fitting regularized multinomial regressions of term frequency against the constituency class mix; the coefficients from these regressions are then used to construct a summary measure of the term frequencies relevant to class mix. Taddy (2013) shows that if the underlying relationship is multinomial, this summary measure is a sufficient statistic for how term usage relates to class mix. Finally, we predict the share of working class residents using this summary measure, to obtain a measure on the same scale as class mix.

The resulting measure captures meaningful variation in class rhetoric. The terms most associated with working class constituencies capture economic issues specific to the working class, especially regulation of working hours, coal mining, and the conditions of tenant farmers and agricultural laborers (Table A.12 shows terms with the most positive and negative

coefficients). The election address with the most working class rhetoric, by a Labour and Liberal candidate in the 1900 election, begins "Dear friends and fellow workmen" and ends by summarizing its platform as "1st. The Rights of Labour. 2nd. The Will of the People." The address at the 95th percentile of working class rhetoric ends with a commitment to "promote the interests of the working classes."

In contrast, addresses with low working class rhetoric tend to emphasize foreign and tariff policy and oppose left-wing policies. The lowest-scoring address, by a Conservative candidate in Hampstead, London, criticizes the Liberals for having "assailed, or threatened, the integrity of the United Kingdom, the Church, the rights of property, and the House of Lords" and advocates for "Efficiency and Economy." That at the 5th percentile focuses almost entirely on the Boer War, tariffs, and military issues.

Working class rhetoric maps predictably onto party differences. Labour candidates use more working class rhetoric than Liberal and Conservative candidates, even after accounting for differences in the parliamentary constituencies in which they campaigned (Figure A.4). That said, there is substantial within-party heterogeneity. The address at the 90th percentile is by a Conservative candidate who endorses the eight hour workday and claims "as regards the textile and other trades I should be guided by the feeling of the workers themselves." Similarly, the 10th percentile address is by a Liberal candidate calling for free trade and "Sound finance." Appendix C prints these addresses in full.

7.2 *Empirical analysis*

To study these outcomes, we average relative class distance over the adult male population of each parish.[12] We estimate regressions of the form

$$y_{ct} = \alpha_t + \beta \, \textit{Relative class distance}_{pt} + \mathbf{x}_{pt}\gamma + \delta_{cp} + \varepsilon_{cpt} \tag{6}$$

where the dependent variable is either the share of the vote won by Labour or our measure of working class rhetoric in candidate manifestos in constituency c at election t. α_t is an election fixed effect, $\mathbf{x}_{pt}$ are parish-level controls, δ_{cp} is a fixed effect for each constituency-parish combination, and ε_{cpt} is the error term.

Because our independent variable, *Relative class distance*$_{pt}$, varies at the parish level, but our dependent variables—Labour voting and election address rhetoric—vary at the

[12]Relative class distance, averaged at the parish level, predicts class identity as measured by names. Table A.11 shows this by regressing the parish–year contributions to variation in working-class name scores on relative class distance, revealing a strong positive correlation. This relationship holds even after conditioning on all time-varying parish controls used in specification 6. Moreover, none of these controls predicts class identity based on names, suggesting that our main independent variable captures a key determinant of identity rather than proxying for other obvious confounders.

constituency level, we spatially intersect parishes and parliamentary constituencies and weight observations by the share of the constituency population accounted for by the parish. This weighting gives each constituency equal weight. We cluster standard errors by parish and constituency to account for dependency across parishes nested in the same constituency and between constituencies containing the same split parishes.[13]

Table 4 shows our estimates of the relationship between relative class distance and voting. Columns (1)–(7) restrict the sample to constituencies with contested elections. As relative class distance increases—bringing the average resident culturally closer to the upper class relative to the working class—support for the Labour Party decreases. A 0.1 unit increase in relative class distance corresponds to an 8 percentage point decrease in the Labour vote, in model (1). Subsequent models add controls for the working class share of the constituency (2), the shares employed in different economic sectors (3), net migration and population growth (4) and the share of Irish-born residents (5). The coefficient on relative class distance is stable and negative across these specifications, and only meaningfully attenuates when adding migration-related controls. These partially capture the same concept as relative class distance. For example, Irish immigrants were culturally distant and tended to be working class. An influx of Irish immigrants would increase the relative distance of residents to the working class relative to the upper class, and so we would expect a similarly negative relationship between Irish immigrants and voting for Labour, as between relative class distance and voting Labour. We find this to be the case.

In the data, a two-standard deviation increase in relative class distance, calculated with the same parish-constituency weights as the regressions, is 0.07. The most restrictive specification thus implies that a two-standard deviation increase in relative class distance corresponds to around a three percentage point decrease in the Labour vote share. This estimate is large; average Labour vote share is less than 5 percentage points. We do not find this decrease in Labour's vote share to clearly benefit either the Conservatives or the Liberals, though the effect on the Liberal share is larger in magnitude (models (6) and (7)).[14] In models (8) and (9), we find that much of the estimated change in Labour's vote share is attributable to the Party's strategic decision over which seats to contest. The coefficient in model (9) indicates that a two–standard deviation increase in relative class distance corresponds to a six–percentage–point decrease in the probability that Labour contested the constituency.

Figure 6 provides evidence that constituencies experiencing increases in relative class

[13]Note that if we aggregated the parish data to the constituency level, we would not be able to account for this dependence across constituencies.

[14]This speaks against a decline in working class identity being identical to an increase in national identity. Were this the case, the Conservative Party, with its emphasis on imperial unity and on symbols of Britishness such as the Crown and the established Church, would have been favored.

	Vote share: Labour					Cons.	Lib.	Labour present	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Relative class distance	-0.797^{**}	-0.792^{**}	-0.716^{**}	-0.576^{**}	-0.439^{*}	0.106	0.283	-1.581^{**}	-0.890^{*}
	(0.239)	(0.241)	(0.222)	(0.220)	(0.211)	(0.140)	(0.218)	(0.426)	(0.383)
Share born in Ireland					-1.362^{*}	0.513^{**}	$1.005^{\dagger}$		-3.170^{**}
					(0.554)	(0.152)	(0.520)		(1.185)
Fixed efffects: Election and constituency-by-parish	x	x	x	x	x	x	x	x	x
Controls: % working class		x	x	x	x	x	x		x
- Sector shares			x	x	x	x	x		x
- Net migration				x	x	x	x		x
Contested elections only	x	x	x	x	x	x	x		
N	37885	37792	37792	37792	37792	37792	37792	46144	46029
R^2	0.624	0.624	0.630	0.635	0.637	0.801	0.644	0.556	0.570

This table presents evidence of the relationship between relative distance to the working class and voting. An observation is a parish-constituency combination, for a given election: 1892, 1900, and the two 1910 elections. In models (1)–(5) the dependent variable is the share of the vote won by Labour in the constituency, in (6) Conservative vote share, in (7) Liberal vote share, in (8)–(9) an indicator that Labour contested the election. Relative class proximity is averaged over the adult male population. Models (2)–(7) and (9) control for the share of adult men employed in working class jobs relative to upper or middle class jobs, (3)–(7) and (9) control for the shares of adult men employed in farming, mining, manufacturing, and services, (4)–(7) and (9) control for population growth with log population and for net migration with the log ratio of adult men resident in the county divided by adult men born in the county. (5)–(7) and (9) control for the share of the population born in Ireland. Observations are weighted by the parish's share of the constituency population in 1891. Models (1)–(7) are restricted to contested elections, (8)–(9) include uncontested elections. Standard errors clustered by parish and constituency in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table 4: Relationship between class distance and voting for Labour

distance were not on different pre-treatment trajectories in voting patterns. Similar to a parallel-trends check, it shows that the change in relative class distance between 1891 and 1911 does not predict either the level of party vote shares in the previous elections of 1885 and 1886 or the change in voting between these two elections. After 1891, however, higher relative class distance is associated with a decrease in the Labour vote share and a corresponding increase in the Liberal vote share.

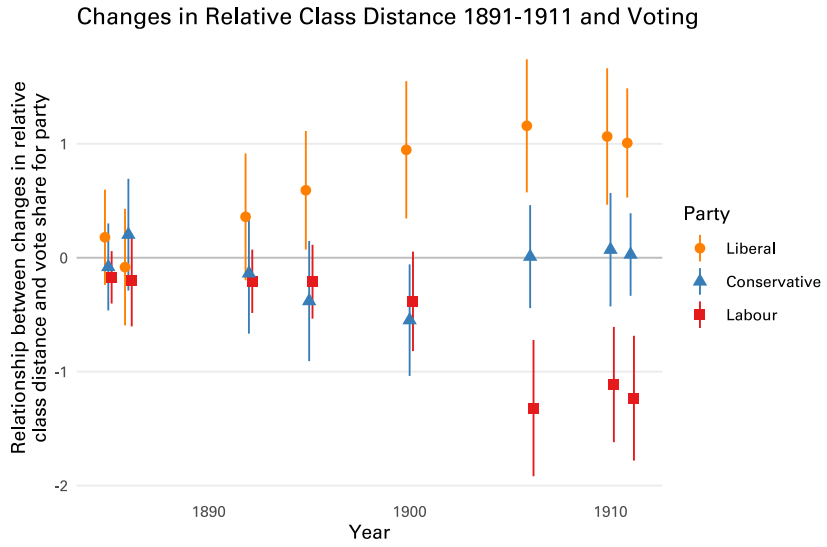


Figure 6: Changes in relative class distance 1891–1911 are uncorrelated with levels and changes in voting pre-1891

This figure shows the relationship between the change in relative class distance between 1891 and 1911, and vote share for different parties. The orange circles are point estimates with 95% confidence intervals from a regression of Liberal Party vote share against the 1891–1911 change in relative class distance, interacted with election indicators, with election fixed effects. Blue triangles are the equivalent for the Conservative vote, red squares for Labour. Each party's coefficients are from a separate regression, with standard errors clustered by parish and constituency.

Candidates campaigning in areas with greater cultural distance from the working class tended to de-prioritize working class issues. Table 5 reports the results of regressions of working class rhetoric against relative class distance. Model (1) indicates that a two standard deviation increase in relative class distance is associated with around a 1 percentage point decrease in working class rhetoric. To place that estimate in perspective, the difference between the average Conservative and Labour election addresses is 2.4 percentage points. (2) shows that this coefficient is largely unchanged controlling for the share of working class residents. (3) interacts the constituency-by-parish fixed effects with indicators for the candidate's party, and so accounts for differences in the mix of parties contesting each constituency. Doing so results in a smaller coefficient, suggesting that the results in models (1) and (2) are partly due to strategic entry, for instance Labour not entering in constituencies with high relative class

	Working class rhetoric					
	(1)	(2)	(3)	(4)	(5)	(6)
Relative class distance	-0.145^{**}	-0.130^{**}	-0.106^{**}	-0.090^{**}	-0.082^{*}	-0.067^{*}
	(0.032)	(0.033)	(0.034)	(0.034)	(0.033)	(0.032)
Share born in Ireland						-0.171^{**}
						(0.058)
Fixed efffects: Election	x	x	x	x	x	x
- Constituency-by-parish	x	x				
- Constituency-by-parish-by-party			x	x	x	x
Controls: % working class		x	x	x	x	x
- Sector shares				x	x	x
- Net migration					x	x
N	161997	161614	161614	161614	161614	161614
R^2	0.725	0.727	0.790	0.790	0.790	0.790

This table presents evidence of the relationship between relative distance from the working class and working class rhetoric, a measure of parliamentary candidates using terminology slanted towards working class voters. An observation is a parish-constituency-candidate combination, for a given election: 1892, 1900, and the two 1910 elections. The dependent variable is working class rhetoric, the extent to which the candidate's rhetoric in their election address is specific to working-class constituencies, calculated using Multinomial Inverse Regression. Relative class proximity is averaged over the adult male population. Models (2)–(6) control for the share of adult men employed in working class jobs relative to upper or middle class jobs, (3)–(6) include fixed effects for the candidate's party interacted with the constituency-parish combination, (4)–(6) control for the shares of adult men employed in farming, mining, manufacturing, and services, (5)–(6) control for population growth with log population and for net migration with the log ratio of adult men resident in the county divided by adult men born in the county. (6) controls for the share of the population born in Ireland. Observations are weighted by the parish's share of the constituency population in 1891. Standard errors clustered by parish and constituency in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table 5: Relationship between class distance and working class rhetoric

distance. Nonetheless, there is a clear negative within-party and constituency relationship, suggesting that in addition to influencing strategic entry, relative class distance influences the positioning of candidates conditional on entry. (4)–(6) add the full set of controls from Table 4; as in Table 4 the main change is from controlling for the share born in Ireland, which is also negatively correlated with working class rhetoric. One might be concerned that these results reflect some idiosyncracy of the multinomial inverse regression process used to generate working class rhetoric scores. To allay that concern, Table A.14 shows that the results of Table 5 are robust to using an alternative measure of working class rhetoric, in which we predict the class mix of the constituency using a Lasso.

7.3 *Why and where did Labour rise?*

We identify relative cultural distance from the working class as a key inhibitor of working-class identification and class-based political mobilization in the period we study. The converse mechanism suggests that class politics should have been favored in locations where workers perceived their class as more culturally similar. How can the patterns of demographic change and shifts in relative class distance during the period help explain the overall rise in Labour support and its spatial distribution?

In the left panel of Figure 7, we plot the change in relative class distance between 1891 and 1911 against the initial share of adult men employed in agriculture. During this period, urban areas experienced a decline in relative distance from the working class, while areas more reliant on farming exhibited the opposite pattern. The right panel of Figure 7 shows that the Labour Party gained votes in areas with low initial agricultural employment.

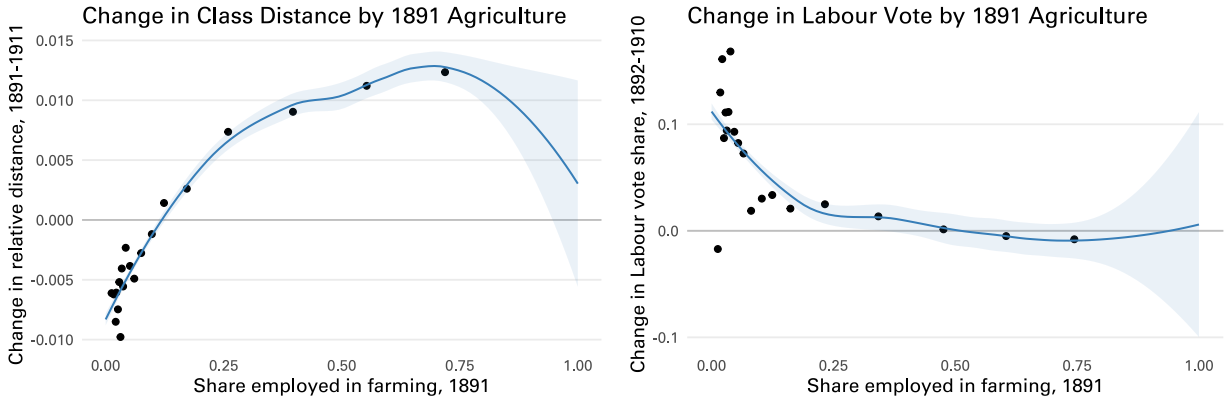


Figure 7: Urban areas experience declining relative class distance and relative increases in Labour voting

The left panel plots the change in relative class distance between 1891 and 1911 against the share of the adult male population employed in farming in 1891, weighted by 1891 population. The right panel plots the change in Labour vote share between the 1892 and December 1910 general elections against 1891 farming employment, weighting parishes by the share of the constituency's 1891 population accounted for by the parish. Black dots are binned averages, blue lines are loess curves with 95% confidence intervals.

Why did relative class distance decrease in cities? This pattern was driven entirely by a drop in the cultural distance of the working class (Figure A.5), which in turn reflected an increase in the share of locally born working-class men (Figure D.1). In Section D of the Appendix, we show that cities experienced a decline in net migration: while these areas attracted distant migrants in 1881, by 1911 the share of locally-born workers—children of earlier migrants—was outpacing new inflows, increasing the cultural proximity of the working class.[15] Changes in the countryside, by contrast, were primarily driven by out-migration

[15]Figure A.6 shows that the number of urban-born and urban-resident men increased during the period.

associated with structural transformation, with the largest losses of local population occurring through occupational shifts out of farming and into the secondary and tertiary sectors (Figures D.3 and D.4).

These patterns suggest an explanation for the time lag in the rise of class-based politics in Britain relative to the timing of industrialization and the growth of the urban proletariat highlighted by much historical literature (Hobsbawm, 1954; Pelling, 1965; Stedman Jones, 1983). Structural transformation drove labor migration into industrial centers in the early stages of industrialization, increasing cultural distance among workers and diluting class identity. As growth in the urban population outpaced in-migration over time, declining relative class distance enabled the formation of shared class identities. Alongside rising support for class-based politics in cities, continued local population loss in the countryside was associated with weaker class identities, contributing to both class and spatial political cleavages.

8 DISCUSSION AND CONCLUSION

This paper introduces an explanation for why and when people identify with their class. Exploiting within-household variation over time, we show that cultural distance from the working class relative to the middle and upper class at the local level reduces working class identification. Through this mechanism, cultural distance influences union membership, voting, and the strategic decisions of politicians to campaign on class lines. In our analysis, migration and demographic change serves to shift relative class distance. Our results therefore connect migration—and structural economic changes that drive internal migration—to a range of political outcomes.

This study is not the first to link diversity to the failure of left-wing movements. Our focus on identity, and direct evidence of cultural distance influencing identity, is distinct from existing accounts. These stress other ways in which diversity undermines class solidarity, including coordination failures and preference heterogeneity that reduce support for collective projects (Alesina and La Ferrara, 2005; Habyarimana et al., 2007).

The historical case has direct relevance for contemporary debates on migration and politics. England and Wales between 1880 and 1910 provide a conservative test of the role of cultural distance: if migration weakened class identity in a period of heightened class salience, driven by industrialization and distributive conflict, then larger effects can be expected today. Moreover, our focus on internal migration understates the degree of heterogeneity we capture. In modern settings—where migration flows are more distant and rising forced migration can

As a result, and despite rising rural to urban migration, the share of the urban population that was born in rural areas declined.

exacerbate cultural mismatch—the effects of cultural distance on social identification and political outcomes are likely to be even greater.

Abrajano, Marisa and Zoltan L. Hajnal. 2015. *White Backlash: Immigration, Race, and American Politics*. Princeton, NJ: Princeton University Press.

Abramitzky, Ran, Leah Boustan and Katherine Eriksson. 2020. "Do immigrants assimilate more slowly today than in the past?" *American Economic Review: Insights* 2(1):125–141.

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2014. "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration." *Journal of Political Economy* 122(3):467–506.

Abramitzky, Ran, Roy Mill and Santiago Perez. 2020. "Linking individuals across historical sources: A fully automated approach." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 53(2):94–111.

Acemoglu, Daron and James A. Robinson. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics* 115(4):1167–1199.

Acemoglu, Daron and James A. Robinson. 2006. "Economic Backwardness in Political Perspective." *American Political Science Review* 100(1):115–131.

Adelman, Paul. 1996. *The Rise of the Labour Party, 1880–1945*. 3rd ed. London: Longman.

Alesina, Alberto, Armando Miano and Stefanie Stantcheva. 2023. "Immigration and redistribution." *The Review of Economic Studies* 90(1):1–39.

Alesina, Alberto, Edward L. Glaeser and Bruce Sacerdote. 2004. *Fighting Poverty in the US and Europe: A World of Difference*. Oxford: Oxford University Press.

Alesina, Alberto and Eliana La Ferrara. 2000. "Participation in heterogeneous communities." *The quarterly journal of economics* 115(3):847–904.

Alesina, Alberto and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43(3):762–800.

Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public goods and ethnic divisions." *The Quarterly journal of economics* 114(4):1243–1284.

Bartolini, Stefano. 2000. *The Political Mobilization of the European Left, 1860–1980: The Class Cleavage*. Cambridge: Cambridge University Press.

Bartolini, Stefano and Peter Mair. 1990. *Identity, Competition, and Electoral Availability: The Stabilisation of European Electorates 1885-1985*. Cambridge: Cambridge University Press.

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.

Brodie, Marc. 2004. *The Politics of the Poor: The East End of London 1885–1914*. Oxford: Clarendom Press.

Bronner, Laura. 2019. Competition and Communication: The Development of Campaigning in Britain from the Second Reform Act to the First World War Phd thesis London School of Economics and Political Science.
URL: *https://etheses.lse.ac.uk/3776/1/Bronner_competition-and-communication.pdf*

Bronner, Laura and Daniel Ziblatt. 2018. The Rise of the Modern Electoral Campaign: the Peers, the People, and the Polarization of the British Party System, 1892–1910. In *Competition and Communication: The Development of Campaigning in Britain from the Second Reform Act to the First World War*. phd thesis II. Identification Number: 10.21953/lse.n0gus1wcchoa.

Cannadine, David. 1998. *Class in Britain*. New Haven: Yale University Press.

Cummins, Neil J and Cormac Ó Gráda. 2025. "The Irish in England." *The Journal of Economic History* 85(1):180–214.

Dewan, Torun, Jaakko Meriläinen and Janne Tukiainen. 2020. "Victorian voting: The origins of party orientation and class alignment." *American Journal of Political Science* 64(4):869–886.

Enamorado, Ted, Benjamin Fifield and Kosuke Imai. 2019. "Using a Probabilistic Model to Assist Merging of Large-Scale Administrative Records." *The American Political Science Review* 113(2):353–371.

Feldman, David. 2000. Migration. In *The Cambridge Urban History of Britain, Volume III: 1840–1950*, ed. Martin Daunton. Cambridge: Cambridge University Press pp. 185–206.

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

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

Fouka, Vasiliki and Marco Tabellini. 2022. "Changing in-group boundaries: The effect of immigration on race relations in the United States." *American Political Science Review* 116(3):968–984.

Fouka, Vasiliki, Soumyajit Mazumder and Marco Tabellini. 2022. "From immigrants to Americans: Race and assimilation during the Great Migration." *The Review of Economic Studies* 89(2):811–842.

Fouka, Vasiliki and Theo Serlin. 2024. Industry and Identity: The Migration Linkage Between Economic and Cultural Change in 19th Century Britain. Technical report NBER Working Paper No. 33114.

Fresh, Adriane. 2024. Population and Political Change in Industrial Britain. Technical report SSRN.
URL: *https://dx.doi.org/10.2139/ssrn.4734362*

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

Gamalerio, Matteo, Mario Luca, Alessio Romarri and Max Viskanic. 2023. "Refugee reception, extreme-right voting, and compositional amenities: Evidence from Italian municipalities." *Regional Science and Urban Economics* 100:103892.

Gilens, Martin. 1999. *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy*. Chicago: University of Chicago Press.

Habyarimana, James, Macartan Humphreys, Daniel N Posner and Jeremy M Weinstein. 2007. "Why does ethnic diversity undermine public goods provision?" *American political science review* 101(4):709–725.

Hessami, Zohal and Sebastian Schirner. 2024. Immigrant Shocks and Shifting Social Group Boundaries. Technical report Working Paper.

History of Parliament Trust and Michael Rush. 2020. "Members of Parliament After 1832 [database].".
URL: *https://membersafter1832.historyofparliamentonline.org*

Hobsbawm, Eric J. 1954. "The Labour Aristocracy in Nineteenth-Century Britain." *The Economic History Review* 7(3):354–376.

Hobsbawm, Eric J. 1984*a*. The Formation of British Working-Class Culture. In *Workers: Worlds of Labor*. New York: Pantheon Books pp. 176–193.

Hobsbawm, Eric J. 1984*b*. The Making of the Working Class 1870-1914. In *Workers: Worlds of Labor*. New York: Pantheon Books pp. 194–213.

Homans, George C. 1969. "The explanation of English regional differences." *Past & Present* (42):18–34.

Inglehart, Ronald. 1977. *The Silent Revolution: Changing Values and Political Styles Among Western Publics*. Princeton: Princeton University Press.

Jusko, Karen Long. 2017. *Who Speaks for the Poor? Electoral Geography, Party Entry, and Representation*. Cambridge Studies in Comparative Politics Cambridge: Cambridge University Press.

Kalyvas, Stathis N. 1996. *The Rise of Christian Democracy in Europe*. Ithaca: Cornell University Press.

Kitschelt, Herbert. 1994. *The Transformation of European Social Democracy*. Cambridge: Cambridge University Press.

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

Lee, Woojin, John Roemer and Karine Van der Straeten. 2006. "Racism, xenophobia, and redistribution." *Journal of the European Economic Association* 4(2-3):446–454.

Lipset, Seymour Martin and Gary Marks. 2000. *It Didn't Happen Here: Why Socialism Failed in the United States*. New York: W. W. Norton & Company.

Lipset, Seymour Martin and Stein Rokkan, eds. 1967. *Party Systems and Voter Alignments: Cross-National Perspectives*. New York: Free Press.

Lonsky, Jakub. 2021. "Does immigration decrease far-right popularity? Evidence from Finnish municipalities." *Journal of Population Economics* 34(1):97–139.

Luttmer, Erzo FP. 2001. "Group loyalty and the taste for redistribution." *Journal of political Economy* 109(3):500–528.

McKibbin, Ross. 1990. *The Ideologies of Class: Social Relations in Britain, 1880–1950*. Oxford: Clarendon Press.

McKibbin, Ross. 1998. *Classes and Cultures: England 1918–1951*. Oxford: Oxford University Press.

Meltzer, Allan H. and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(5):914–927.

Miguel, Edward and Mary Kay Gugerty. 2005. "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of public Economics* 89(11-12):2325–2368.

Pelling, Henry. 1965. *A Short History of the Labour Party*. London: Macmillan.

Pelling, Henry. 1967. *Social Geography of British Elections, 1885–1910*. London: Macmillan.

Phythian-Adams, Charles. 1993. "Local history and societal history." *Local Population Studies* 51(Autumn):30–45.

Przeworski, Adam and John Sprague. 1986. *Paper Stones: A History of Electoral Socialism*. Chicago: University of Chicago Press.

Roemer, John E. 1998. "Why the poor do not expropriate the rich: an old argument in new garb." *Journal of Public Economics* 70(3):399–424.

Rush, Michael. 2001. *The Role of the Members of Parliament Since 1868: From Gentlemen to Players*. New York: Oxford University Press.

Schürer, Kevin and Edward Higgs. 2024. "Integrated Census Microdata (I-CeM), 1851-1911.". SN: 7481.
URL: *http://doi.org/10.5255/UKDA-SN-7481-3*

Seventy-fourth Annual Report of the Registrar-General of Births, Deaths, and Marriages, in England and Wales (1911). 1913. London: H.M. Stationery Office, Darling and Son Ltd. Cd. 6578.

Shayo, Moses. 2009. "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution." *American Political Science Review* 103(2):147–174.

Shayo, Moses. 2020. "Social Identity and Economic Policy." *Annual Review of Economics* 12:355–389.

Stedman Jones, Gareth. 1974. "Working-Class Culture and Working-Class Politics in London, 1870–1900: Notes on the Remaking of a Working Class." *Journal of Social History* 7(4):460–508.

Stedman Jones, Gareth. 1983. *Languages of Class: Studies in English Working Class History, 1832–1982*. Cambridge: Cambridge University Press.

Steinmayr, Andreas. 2021. "Contact versus exposure: Refugee presence and voting for the far right." *Review of Economics and Statistics* 103(2):310–327.

Taddy, Matt. 2013. "Multinomial Inverse Regression for Text Analysis." *Journal of the American Statistical Association* 108(503):755–770.

Tajfel, H and JC Turner. 1986. The Social Identity Theory of Inter-Group Behavior. In *Psychology of Intergroup Relations*. Chigago: Nelson-Hall pp. 7–24.

Turner, John C., Michael A. Hogg, Penelope J. Oakes, Stephen D. Reicher and Margaret S. Wetherell. 1987. *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford: Basil Blackwell.

Wald, Kenneth D. 1983. *Crosses on the Ballot: Patterns of British Voter Alignment since 1885*. Princeton NJ: Princeton University Press.

Zhou, Yang-Yang, Guy Grossman and Shuning Ge. 2023. "Inclusive Refugee-Hosting in Uganda Improves Local Development and Prevents Public Backlash." *World Development* 166:106203.

Online Appendix for Migration and the Making of the English Middle Class

Table of Contents

A ADDITIONAL FIGURES AND TABLES

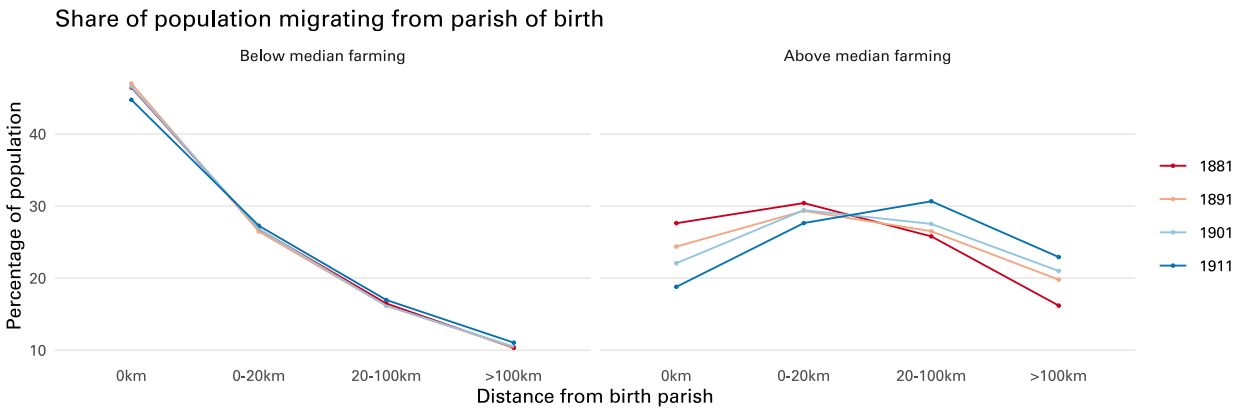


Figure A.1: People migrated over longer distances, especially from rural areas

This figure plots the share of the population living at different distances from their parish of birth, in successive censuses, subset by whether the parish of birth had an above- or below-median share of employment in farming in 1881.

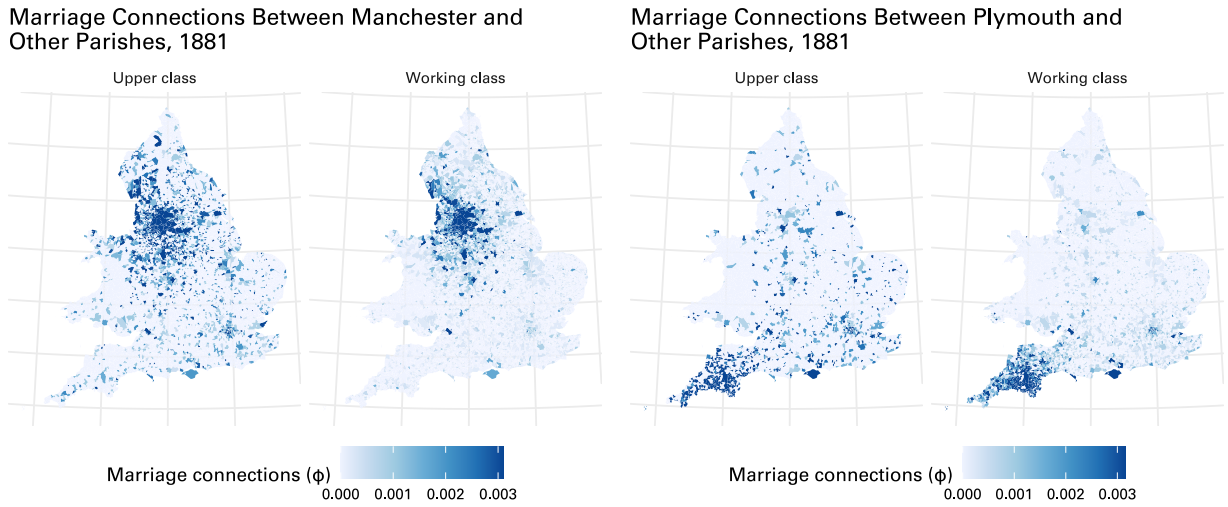


Figure A.2: Marriage connections by class between Manchester and Plymouth and other parishes

This figure shows our measure of marriage connections using data from the 1881 census, between each parish and Manchester (left) or Plymouth (right).

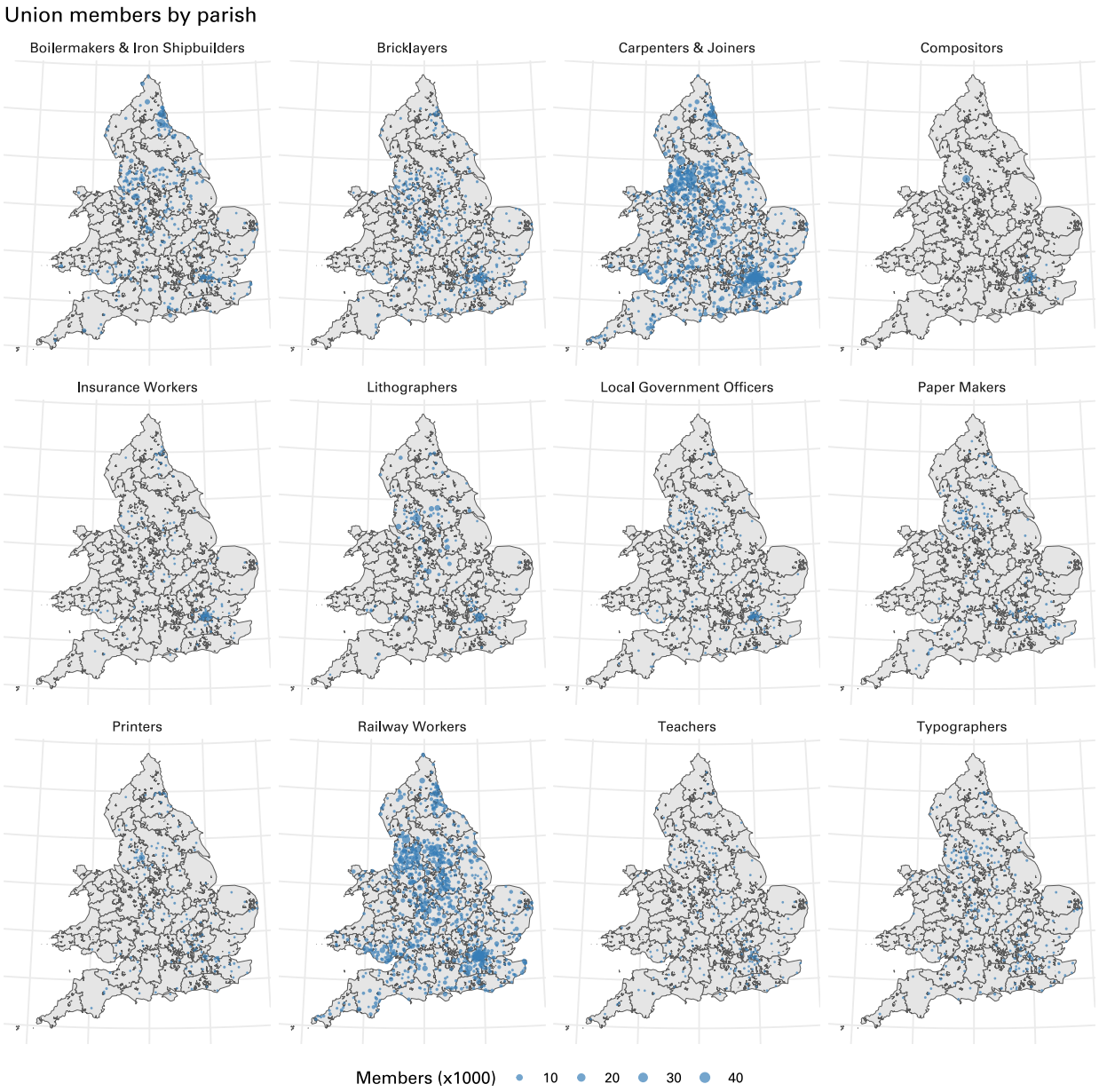


Figure A.3: Union members by parish

Each map shows the spatial distribution of union members for a given union. Each dot shows the number of union members in a given parish in that union over the entire period analyzed, in our dataset of union members (prior to merging this data into the census). The size of the dot corresponds to the number of members.

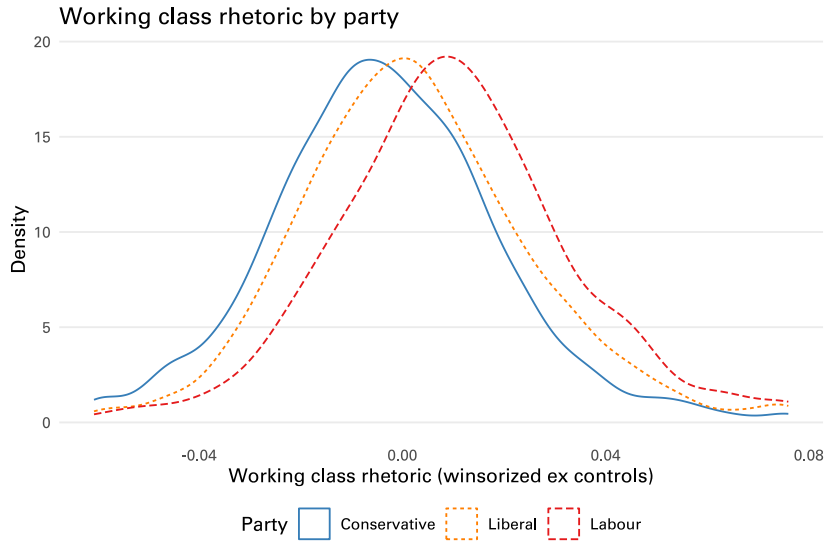


Figure A.4: Labour candidates use more working class rhetoric

This figure shows the density of working class rhetoric, after residualizing out the actual working class percentage of the constituency, and constituency and election fixed effects, by party. The dashed red line is for Labour candidates; that it is to the right of the dotted orange (Liberal) and solid blue (Conservative) lines indicates that Labour candidates, even adjusting for constituency features, used more working class rhetoric. To improve legibility, residualized working class rhetoric is winsorized at the 0.5th and 99.5th percentiles.

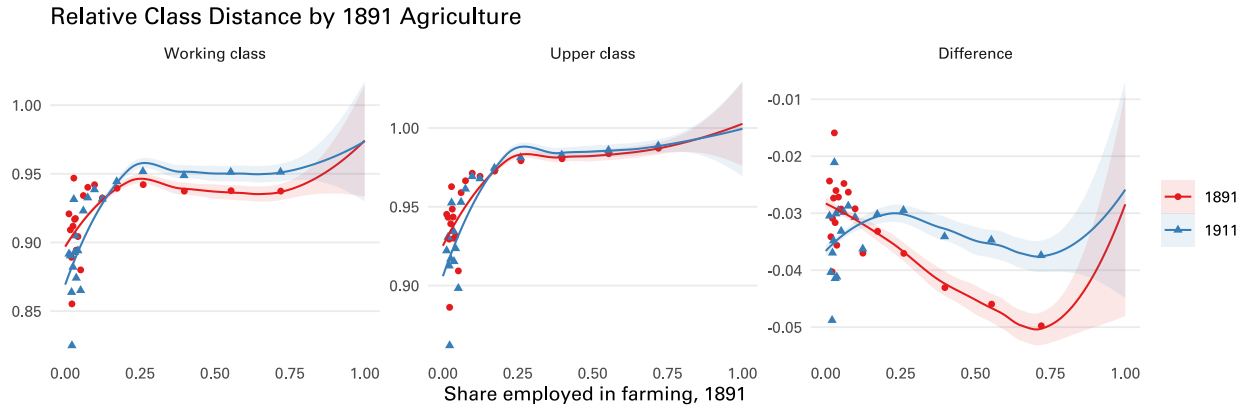


Figure A.5: Rising relative distance to the working class in rural areas is driven by an increase in absolute distance to the working class

This figure plots average distance to the working class, the upper class, and the difference between them ("relative class distance") against the share employed in farming in 1891, weighted by 1891 population. Dots and triangles are binned averages, lines are loess curves with 95% confidence intervals.

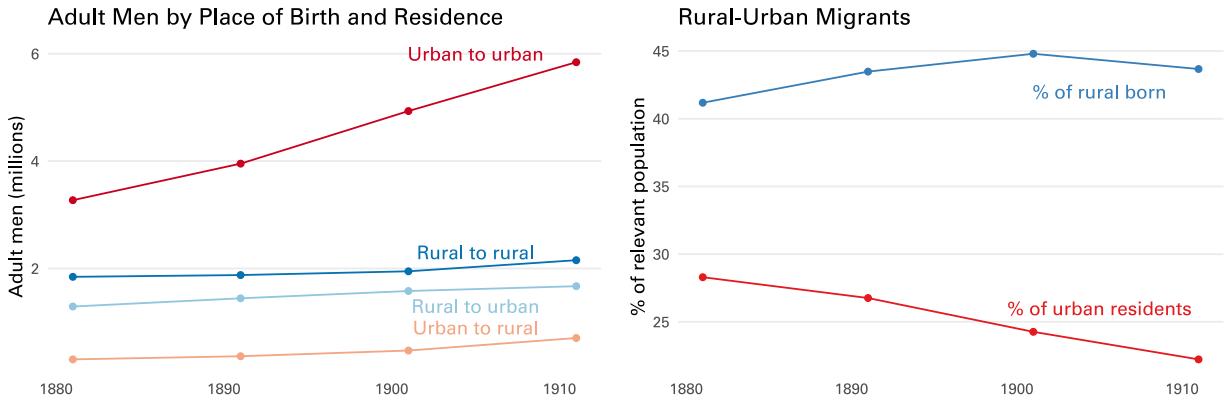


Figure A.6: Rising rural-to-urban migration, offset by a growing urban-born population

The left figure plots the share of adult men born in urban or rural areas, living in urban or rural areas. The right figure plots rural to urban migrants as a percentage of the rural-born and urban-resident populations. Areas are classified as rural or urban depending on whether they have a population density of greater than 300 persons per square kilometer in 1881.

Table A.1: Most and least working class names appearing 100 or more times

Most working class					Least working class				
Metaphone	Name	Sex	n	Score	Metaphone	Name	Sex	n	Score
PTS	Patsy	M	284	84.7	SRL	Cyril	M	193	4.4
PT	Pat	M	795	82.2	RJNLT	Reginald	M	1157	5.2
MKL	Mickel	M	201	79.5	SSL	Cecil	M	605	5.7
BRJT	Bridget	F	12053	79.3	JRLT	Gerald	M	402	6.8
PTRK	Patrick	M	29906	78.9	ALJRNN	Algernon	M	329	7.0
XTRK	Shadrack	M	367	77.9	ESTS	Eustace	M	215	7.0
TLL	Delilah	F	138	77.4	MKS	Max	M	255	7.9
SXN	Sushannah	F	434	76.8	KLT	Claude	M	452	8.8
RSHN	Rosehannah	F	224	76.0	KRTN	Gordon	M	351	9.1
ES	Esau	M	577	76.0	HRLT	Harold	M	1125	9.1
N	Noah	M	2711	74.6	KRHM	Graham	M	167	9.3
JB	Job	M	6692	73.2	KRNFL	Granville	M	117	9.3
ABRM	Abram	M	1354	73.1	KNSTNS	Constance	F	595	9.7
ENK	Enock	M	660	72.9	MNTK	Montague	M	461	9.7
KMFRT	Comfort	F	191	72.9	TTL	Dudley	M	155	9.8
NMRT	Nimrod	M	114	72.2	KSTFS	Gustavus	M	176	10.2
SRHN	Sarahann	F	314	72.0	HK	Hugo	M	166	10.4
EL	Eli	M	6193	71.9	LSL	Leslie	M	158	10.7
SKR	Squire	M	1346	71.9	OKTF	Octavia	F	109	10.8
XTRX	Shadrach	M	336	71.9	HWRT	Howard	M	699	11.0

Table A.2: Occupations with highest average working class scores in 1911 census

Code	Occupation	n	Score
408	builders' excavators	199	48.7
588	sacking manufacture (various)	854	47.9
582	jute manufacture (various)	628	47.9
548	cotton & cotton goods manufacture card blowing room processes	40890	47.5
601	knitter (undefined)	559	47.3
551	cotton & cotton goods manufacture weaving processes	148957	47.1
550	cotton & cotton goods manufacture winding warping processes	55902	47.0
605	weavers sundry fabrics	2293	46.9
575	crepe gauze manufacture (various)	129	46.9
606	weavers (undefined)	52181	46.8
466	earthenware figure and image makers	1184	46.7
227	slate quarriers	7712	46.7
556	fustian manufacture	3126	46.6
744	rag dealers	6850	46.3
354	shipyard labourers (undefined)	6947	46.2
562	worsted and stuff manufacture winders warpers weavers	8747	46.2
580	flax linen and damask manufacture (various)	2099	46.1
563	flannel manufacture (various)	371	46.1
432	railway labourers navvies (contractors labourers) default	13016	46.1
356	others in ship/boat building - metal	151	46.1

Table A.3: Occupations with lowest average working class scores in 1911 census

Code	Occupation	n	Score
26	royal marines (retired officers)	104	39.1
23	navy officers retired and coastguards rn reserve officers	944	39.3
69	surveyors assistant	833	39.5
121	bank service, bank officials and clerks	26146	39.6
71	sculptors (artists)	628	39.7
281	railway - signal switch turntable makers	130	39.8
429	railway contractors	507	39.8
123	insurance officials clerks	24507	39.9
781	medical students	3209	39.9
72	architects	6167	40.1
786	agricultural students	1893	40.2
6	senior government officials, judges, public lawyers and civil service	20341	40.2
19	army officers retired	5277	40.2
634	silk merchants dealers	384	40.3
115	accountants	12484	40.3
8	senior officials and others east india service	4955	40.3
60	others scientific	4224	40.3
120	bankers	1781	40.3
64	mining engineers	1119	40.3
38	barristers (private) advocate	2772	40.4

	Child's working class score			
	(1)	(2)	(3)	(4)
Working class occupation	2.225**	2.130**		
	(0.024)	(0.024)		
Occupational status			−0.090**	−0.084**
			(0.001)	(0.001)
Fixed effects: Father's name		x		x
N	2610864	2603156	2138647	2131724
R^2	0.004	0.010	0.005	0.011

This table presents evidence of the relationship between the occupations of fathers in the 1911 census and the average working class score of their children. In models (1) and (2), the independent variable is an indicator that the father has an occupation classified as working class rather than upper or middle class by the Registrar General, in (3) and (4) the HISCAM occupational status score for the father's occupation. (2) and (4) include fixed effects for the father's first name. Robust standard errors in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.4: Relationship between fathers' class and children's names

	Log marriage rate: Upper class					Working class				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log geographic distance	−0.191**	−0.171**	−0.170**	−0.167**	−0.162**	−0.415**	−0.385**	−0.383**	−0.368**	−0.359**
	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Same parish	−0.191**	−0.031	−0.021	−0.004	−0.206*	−1.071**	−0.820**	−0.804**	−0.703**	−1.201**
	(0.073)	(0.062)	(0.062)	(0.065)	(0.091)	(0.049)	(0.044)	(0.043)	(0.045)	(0.078)
Same county		0.081**	0.080**	0.080**	0.082**		0.119**	0.116**	0.110**	0.112**
		(0.010)	(0.010)	(0.010)	(0.010)		(0.005)	(0.005)	(0.006)	(0.006)
Same country			0.051**	0.040**	−0.102**			0.129**	0.067**	−0.124**
			(0.014)	(0.014)	(0.031)			(0.018)	(0.018)	(0.024)
Religious distance				−0.049*	−0.013				−0.294**	−0.258**
				(0.021)	(0.022)				(0.016)	(0.017)
Surname distance					−1.617**					−2.676**
					(0.270)					(0.217)
Fixed efffects: Origin parish	x	x	x	x	x	x	x	x	x	x
- Destination parish	x	x	x	x	x	x	x	x	x	x
N	240466	240466	240466	240275	235655	1336440	1336440	1336440	1334999	1301237
R^2	0.873	0.873	0.874	0.873	0.873	0.766	0.767	0.767	0.768	0.769

This table shows predictors of marriage rates between counties, φ_{il}^{j}, calculated as the number of people born in a given parish married to people born in another parish, normalized by the geometric mean of the number of married people born in both parishes. Data is at the level of pairs of parishes. In models (1)–(5), the dependent variable is the log marriage rate for upper class couples. In (6)–(10) the equivalent for working class couples, both using data from the 1881 census. All models include fixed effects for both sets of parishes. (1) and (6) include coefficients for the log distance between parishes and an indicator that the two parishes are the same, (2) and (7) add an indicator that both parishes are in the same county, (3) and (8) add an indicator that both parishes are in the same country (i.e. Wales or England), (4) and (9) add the Euclidean distance of religious attendance shares in the 1851 census, (5) and (10) add the Euclidean distance of surname shares among adult men in the 1861 census. Standard errors clustered by both sets of parishes in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.5: Correlates of marriage rates between parishes

	Working class score	
	(1)	(2)
Union member	0.753**	0.952**
	(0.021)	(0.024)
Fixed efffects: Year	x	
- Occupation × year		x
N	1760932	1760929
R^2	0.037	0.064

This table presents evidence of the relationship between union membership and the working class score given to children. Data is at the individual-by-census year level, restricted to those in occupations for which we have union membership data. The dependent variable is the working class score of the individual's children. Model (1) includes census year fixed effects, (2) interacts those with fixed effects for the individual's occupation. Robust standard errors in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.6: Relationship between union membership and children's working class scores

	Union member							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance to occupation	$0.052^{\dagger}$	-0.033^{**}	-0.044^{**}	-0.022^{**}	0.302^{**}	-0.003	0.001	0.010
	(0.028)	(0.006)	(0.006)	(0.005)	(0.048)	(0.009)	(0.010)	(0.010)
Distance to working class					-0.302^{**}	-0.023^{**}	-0.034^{**}	-0.024^{**}
					(0.042)	(0.008)	(0.008)	(0.008)
Fixed efffects: Occupation × year	x				x			
- Occupation × parish × year		x	x	x		x	x	x
- Occupation × birth parish × year			x	x			x	x
Controls: Gender, age				x				x
N	6039038	5917587	5698799	5684225	6039038	5917587	5698799	5684225
R^2	0.281	0.409	0.465	0.467	0.287	0.409	0.465	0.467

This table presents evidence of the relationship between distance to other members of one's occupation and union membership. Each observation is an individual worker in a given census year, for the period 1881–1911. The sample is restricted to those employed in occupations related to the Carpenters and Joiners, Railway Workers, Teachers, Boilermakers and Iron Shipbuilders, Typographers, Lithographers, Bricklayers, Papermakers, Printers, Local Government Officers, Insurance Workers, and Compositors unions. The dependent variable takes a value of 1 if the worker is a member of one of these unions, 0 otherwise. The independent variable of interest is average distance based on birthplaces between the individual and others in the same parish employed in the same occupation. Models (1) and (5) include occupation-year fixed effects, (2)–(4) and (6)–(8) occupation-parish-year fixed effects, which subsume occupation-year fixed effects, (3)–(4) and (7)–(8) add occupation-birth parish-year fixed effects, (4) and (8) also control for gender and a third-degree polynomial in age. (5)–(8) control for cultural distance to the working class residents of the parish. OLS estimates. Standard errors clustered by parish and birth parish in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.7: Relationship between class distance within occupation and union membership

	Union member					
	(1)	(2)	(3)	(4)	(5)	(6)
Relative class distance	−0.050	−0.063**	−0.072**	−0.034**	−0.040**	0.003
	(0.041)	(0.010)	(0.009)	(0.008)	(0.010)	(0.003)
Fixed efffects: Occupation × year	x					
- Occupation × parish × year		x	x	x	x	x
- Occupation × birth parish × year			x	x	x	x
Controls: Gender, age				x	x	x
Sample:	All	All	All	All	Working	Upper
N	6052902	5930345	5707463	5692870	4951251	688950
R^2	0.250	0.384	0.443	0.445	0.434	0.288

This table replicates Table 3, using a deterministic algorithm to merge union membership data into the census, in place of the probabilistic algorithm used there. Our algorithm iterates through combinations of first name, initials, last name, metaphone codes for first or last name, parish, and county, requiring an exact match on all variables between the census and union membership data, in all cases also requiring the census observation's occupation to match the union sector of employment. The table presents evidence of the relationship between relative distance to the working class and union membership. Each observation is an individual worker in a given census year, for the period 1881–1911. The sample is restricted to those employed in occupations related to the Carpenters and Joiners, Railway Workers, Teachers, Boilermakers and Iron Shipbuilders, Typographers, Lithographers, Bricklayers, Papermakers, Printers, Local Government Officers, Insurance Workers, and Compositors unions. The dependent variable takes a value of 1 if the worker is a member of one of these unions, 0 otherwise. If n census observations in a given year merge to the same union member observation, the dependent variable is $1/n$ for each of these census observations; the dependent variable can be thought of as the probability that the individual is a union member. Model (1) includes occupation-year fixed effects, (2)–(6) occupation-parish-year fixed effects, which subsume occupation-year fixed effects, (3)–(6) add occupation-birth parish-year fixed effects, (4)–(6) also control for gender and a third-degree polynomial in age. In (5) the sample is further restricted to those in occupations classified as working class by the Registrar-General, in (6) to those classified as upper or middle class. OLS estimates. Standard errors clustered by parish and birth parish in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.8: Relationship between class distance and union membership, using a deterministic merging algorithm

	Successful merge				
	(1)	(2)	(3)	(4)	(5)
Working class score	0.000	0.001	0.001	0.001	0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Length of first name		0.047**	0.047**	0.047**	0.047**
		(0.005)	(0.005)	(0.005)	(0.005)
More than one with same first name			0.348**	0.347**	0.340**
			(0.016)	(0.016)	(0.016)
Length of last name				−0.009**	−0.008**
				(0.001)	(0.001)
More than one with same last name					0.212**
					(0.008)
Fixed effects: Union	x	x	x	x	x
Controls: year	x	x	x	x	x
N	1671083	1671083	1671083	1671083	1671083
R^2	0.236	0.277	0.278	0.279	0.281

This table presents evidence of the relationship between each individual's working class score, and the probability of them being successfully merged into the census. The sample is the dataset of union members, each observation is an individual union member. The dependent variable takes a value of 1 if the individual is successfully merged into the census, 0 if not. They independent variable of interest is the working class score for the individual's name. All models include fixed effects for the Standard errors clustered by the first name's metaphone code used to assign upper class scores in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.9: Working class scores are uncorrelated with the probability of a successful merge

	Union member		Deterministic merge	
	(1)	(2)	(3)	(4)
Relative class distance	-0.030^{**}	-0.026^{**}	-0.034^{**}	-0.029^{**}
	(0.009)	(0.009)	(0.008)	(0.008)
Fixed efffects: Occupation × parish × year	x	x	x	x
- Occupation × birth parish × year	x	x	x	x
Controls: Gender, age	x	x	x	x
- Merge success predictors		x		x
N	5692870	5692870	5692870	5692870
R^2	0.467	0.468	0.445	0.448

This table presents evidence of the relationship between relative distance to the working class and union membership, controlling for features of names that predict success in merging observations from the union members dataset into the census. Each observation is an individual worker in a given census year, for the period 1881–1911. The sample is restricted to those employed in occupations related to the Carpenters and Joiners, Railway Workers, Teachers, Boilermakers and Iron Shipbuilders, Typographers, Lithographers, Bricklayers, Papermakers, Printers, Local Government Officers, Insurance Workers, and Compositors unions. In (1)–(2) the dependent variable is union membership as determined by merging the individual into the union members dataset with the main probabilistic algorithm, in (3)–(4) the same but using the deterministic algorithm used in Table A.8. All models include occupation-parish-year and occupation-birth parish-year fixed effects and controls for gender and a third-degree polynomial in age. (2) and (4) also control for features of names that predict merge success: the length of the first name and surname, and an indicator that the first and surname appears more than once in the union members dataset. OLS estimates. Standard errors clustered by parish and birth parish in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.10: Relationship between class distance and union membership, robustness to controlling for predictors of merge success

	Parish-year class fixed effect				
	(1)	(2)	(3)	(4)	(5)
Relative class distance	-6.744^{**}	-6.770^{**}	-5.783^{**}	-4.602^{*}	-3.987^{*}
	(1.691)	(1.718)	(1.719)	(1.855)	(1.825)
Share working class		−0.180	−0.906	−0.777	−0.728
		(1.633)	(1.963)	(2.026)	(2.039)
Share farming			0.971	0.816	0.709
			(0.653)	(0.645)	(0.655)
Share mining			0.961	0.681	0.608
			(1.051)	(1.100)	(1.098)
Share manufacturing			−1.664	$-1.872^{\dagger}$	$-1.801^{\dagger}$
			(1.014)	(1.019)	(1.022)
Share services			−1.012	−1.012	−0.879
			(1.455)	(1.467)	(1.472)
log population				0.294	0.271
				(0.226)	(0.225)
log net migration				−0.297	−0.267
				(0.222)	(0.227)
Share born in Ireland					−6.347
					(4.608)
Fixed efffects: Year and parish	x	x	x	x	x
N	30033	30033	30033	30033	30033
R^2	0.660	0.660	0.661	0.661	0.662

This table presents evidence of the relationship between relative distance to the working class and class identity, at the parish level, using as the dependent variable the parish-year fixed effect from a regression of working class scores against parish-year and household fixed effects. Each observation is a parish-year combination, for 1891, 1901, and 1911. Relative class distance is averaged over the adult male population. All models include fixed effects for the election and constituency-parish unit. Model (2) adds controls for the share of adult men employed in working class class jobs relative to upper or middle class jobs, (3) for the shares of adult men employed in farming, mining, manufacturing, and services, (4) for population growth with log population and for net migration with the log ratio of adult men resident in the county divided by adult men born in the county, (5) for the share of the population born in Ireland. Observations are weighted by the parish's 1891 population. Standard errors clustered by parish in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.11: Relationship between class distance and parish-year class identity fixed effects

Table A.12: Terms most predictive of upper and working-class constituencies, of those used more than 200 times

Most working class		Least working class	
Term	Coefficient	Term	Coefficient
hour bill	12.706	municip	-5.901
mid	12.657	liber radic	-5.740
farmer	10.864	staunch	-3.912
divis counti gentlemen	8.680	graduat	-3.675
counti gentlemen	8.674	elector reform	-3.634
divis counti	8.545	elector gentlemen	-3.438
miner	7.829	ratepay	-3.428
agricultur	7.016	1909	-3.398
coal	6.885	counti council	-3.367
eastern	6.184	train	-3.296
villag	6.138	10	-3.234
royalti	5.143	ancient	-3.203
farm	5.094	dominion	-3.175
hold act	5.061	fix	-3.045
agricultur labour	4.783	ground valu	-3.028
tenant	4.460	wholli	-3.006
agricultur land	4.355	instanc	-2.999
parliamentari divis	4.261	incom tax	-2.950
small hold act	4.245	incom	-2.916
visit	4.199	distribut	-2.903
associ divis	4.161	lord salisburi govern	-2.827
eight hour	4.126	salisburi govern	-2.794
labour parti	4.083	leasehold	-2.794
small hold	4.017	shorten	-2.728
parti divis	3.966	taxat ground	-2.662
chariti	3.964	committe room	-2.574
pound	3.956	local affair	-2.513
elect member	3.950	port	-2.507
liber associ	3.899	grow	-2.420
religi equal	3.841	diminish	-2.371

Table A.13: Terms most predictive of working and upper-class constituencies, Lasso coefficients

Most working class		Least working class	
Term	Coefficient	Term	Coefficient
cabl	4.47	especi food	-40.56
miner	3.49	also urgent	-32.98
counti gentlemen	2.83	crescent	-29.99
agricultur	2.81	nw	-20.31
tax necess	2.14	wc	-20.25
industri district	2.10	radic free trade	-18.01
north east	2.06	re distribut	-17.12
coal	1.97	bill allow	-16.59
great industri	1.81	gener parliamentari elect	-16.38
north west	1.67	divis poll	-16.26
industri constitu	1.65	coastguard	-15.79
royalti	1.52	work found	-15.08
divis	1.44	fix incom	-13.73
small hold	1.33	w decemb	-13.12
form religion	1.29	singular	-13.11
elector parliamentari	1.11	advanc temper	-12.90
fund trade	1.05	r p	-12.49
eight hour bill	0.94	middl work class	-11.43
mid	0.88	believ lord	-11.16
(Intercept)	0.88	parti whether	-10.35
reform poor	0.84	suffer poor	-10.28
wilt	0.68	must still	-9.98
ag	0.62	1906 elector divis	-9.93
joicei	0.54	tramwai	-9.86
farmer	0.50	pm	-9.81
abolit hereditari	0.48	squar w	-9.74
poor poorer	0.45	suburb	-9.55
tori	0.44	countri permit	-9.10
mine	0.43	clerk	-9.04
honour gentlemen faithfulli	0.39	radic associ	-8.89

	Working class rhetoric (Lasso)					
	(1)	(2)	(3)	(4)	(5)	(6)
Relative class distance	-0.056^{**}	-0.051^{**}	-0.050^{**}	-0.041^{**}	-0.033^{*}	-0.033^{*}
	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.014)
Share born in Ireland						0.001
						(0.035)
Fixed efffects: Election	x	x	x	x	x	x
- Constituency-by-parish	x	x				
- Constituency-by-parish-by-party			x	x	x	x
Controls: % working class		x	x	x	x	x
- Sector shares				x	x	x
- Net migration					x	x
N	161997	161614	161614	161614	161614	161614
R^2	0.634	0.635	0.721	0.722	0.722	0.722

This table presents evidence of the relationship between relative distance to the working class and working class rhetoric, a measure of parliamentary candidates using terminology slanted towards working class voters, calculated with Lasso rather than Multinomial Inverse Regression. An observation is a parish-constituency-candidate combination, for a given election: 1892, 1900, and the two 1910 elections. The dependent variable is working class rhetoric, the extent to which the candidate's rhetoric in their election address is specific to working-class constituencies, calculated using Lasso. Relative class proximity is averaged over the adult male population. All models include fixed effects for the election. Models (1)–(2) include fixed effects for the constituency-parish combination. (2)–(6) control for the share of adult men employed in working class jobs relative to upper or middle class jobs, (3)–(6) include fixed effects for the candidate's party interacted with the constituency-parish combination, (4)–(6) control for the shares of adult men employed in farming, mining, manufacturing, and services, (5)–(6) control for population growth with log population and for net migration with the log ratio of adult men resident in the county divided by adult men born in the county. (6) controls for the share of the population born in Ireland. Observations are weighted by the parish's share of the constituency population in 1891. Standard errors clustered by parish and constituency in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table A.14: Relationship between class distance and upper class rhetoric, calculated by Lasso

B ROBUSTNESS OF IDENTITY RESULTS

This appendix details additional results that confirm the robustness of the relationship between relative class distance and class identity.

Our estimates are not sensitive to how we measure cultural proximity. Table B.1 shows that the results of Table 2 are robust to alternative measures of relative class distance based on geographic distance, the similarity of surnames between locations in 1861, and the similarity of religious denominations in 1851. Our estimates are also robust to alternative ways of computing working class scores. Although our preferred measure relies on names of adults in the 1881 census, which predates the naming choices we analyze, one might argue that the class valence of names changes over time. Table B.2 replicates our analysis using name scores computed among adults in the 1901 census, as well as time-varying scores computed separately for each census year analyzed. Estimates remain similar.

One concern is that the class valence of names may have varied across regions. Relatedly, it may capture broad regional and religious differences—for instance, Anglican names being more upper class, and Nonconformist or Irish names being lower class. To address these concerns, we construct location-specific working class scores using information on names held by people in the child's county of birth. We also compute working class scores for *cultural provinces*—groups of counties identified by historian Charles Phythian-Adams as culturally linked due to shared geographic features (Phythian-Adams, 1993). These measures rely on substantially less data for the construction of name scores and are therefore noisier. Nonetheless, Table B.3 shows that our estimates remain robust in both direction and magnitude.

Given the prevalence of distinctively Irish names like Patrick and Bridget among the most working class names in Table A.1, one might be concerned that our estimates are picking up Irish, rather than class, identity. In Table B.4 we investigate this alternative explanation by constructing a measure of how distinctive each name was to Irish immigrants. We find no relationship between relative class distance and distinctively Irish naming, and find that our main results are robust to dropping distinctively-Irish names.

Our identifying variation comes from households with at least two children who were either born across different census periods in parishes that experienced changes in relative class distance, or in different parishes due to household moves between births. To assuage concerns that this may represent an unrepresentative or otherwise atypical sample, we estimate specifications analogous to those in Table 2, but using fixed effects for the father's occupation rather than for the household. In other words, instead of comparing siblings within the same household over time, we now compare changes in name scores across all households within the same occupation. This approach relies on stronger identification assumptions—specifically, that occupations rather than households follow parallel trends in naming—but it also allows us to exploit more data and additional variation, providing evidence that our main findings are not an artefact of the sample used.

Table B.5 shows that this approach yields similarly positive but larger estimates. This is expected: household fixed effects tend to exacerbate measurement error and bias coefficients downward. Columns 9 and 10 restrict the sample to children born in the same parish as their father and exclude households that moved across parishes, effectively implementing a difference-in-differences design at the parish (or parish-occupation) level. In this specification, we estimate that moving from the 5th to the 95th percentile of relative class distance corresponds to a decrease in working class name scores of 0.54—about 28% of the difference between Oxbridge and non-graduate MPs.

Our main findings indicate that greater relative distance from the working class decreases the likelihood that parents give their children working-class names. We interpret this as evidence of weaker identification with the working class. An alternative interpretation is that naming may

	Working class score											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Relative class distance:												
Geographic proximity	−0.328**			−0.292**			−0.222*			−0.292**		
	(0.078)			(0.080)			(0.088)			(0.112)		
Surname similarity		−0.358**			−0.308**			−0.287**			−0.267*	
		(0.079)			(0.083)			(0.098)			(0.113)	
Religious similarity			−0.194**			−0.166**			−0.082			−0.242**
			(0.055)			(0.058)			(0.075)			(0.075)
Fixed efffects: Household	x	x	x	x	x	x	x	x	x	x	x	x
- Year	x	x	x	x	x	x				x	x	x
- Parish				x	x	x						
- Parish × year							x	x	x			
- Parish × father's parish of birth										x	x	x
Controls: Sex, birth order	x	x	x	x	x	x	x	x	x	x	x	x
R^2	0.469	0.469	0.469	0.470	0.470	0.470	0.482	0.482	0.482	0.504	0.504	0.504

This table presents evidence of the relationship between distance to the working or upper class and working class naming, calculating distance using the geographic proximity of birthplaces, the similarity of surname shares among adult men in the 1861 census, and the similarity of religious denomination shares in the 1851 census of religion, rather than inter-marriage between parishes. Data is at the level of births, differentiated by household, year, and parish of birth. The dependent variable is the Working Class Score for the child. All models include household fixed effects and controls for sex and birth order, (1)-(6) and (10)–(12) include birth year fixed effects, (4)–(6) include parish of birth fixed effects, (7)–(9) parish of birth by year of birth fixed effects which subsume birth year and parish fixed effects, and (10)–(12) fixed effects for the parish of birth interacted with the father's parish of birth, which subsume parish fixed effects. Standard errors clustered by household and parish in parentheses. The number of observations in all models is 25,773,095. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.1: Within-father relationship between class distance and upper class naming, calculating relative class distance using geographic proximity, surname similarity, and religious similarity

	Working class score: 1901 data				Decade-specific data			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative class distance	−0.747**	−0.803**	−0.382*	−1.948**	−0.838**	−0.883**	−0.553**	−1.650**
	(0.182)	(0.190)	(0.174)	(0.513)	(0.211)	(0.225)	(0.210)	(0.568)
Fixed efffects: Household	x	x	x	x	x	x	x	x
- Year	x	x		x	x	x		x
- Parish		x				x		
- Parish × year			x				x	
- Parish × father's parish of birth				x				x
Controls: Sex, birth order	x	x	x	x	x	x	x	x
R^2	0.484	0.485	0.497	0.519	0.475	0.475	0.487	0.510

This table presents evidence of the relationship between distance to the working or upper class and working class naming, calculating working class scores using either data from the 1901 census (1–4) or the previous census (5–8), for the decades 1881 and 1891 we use the 1881 census, for 1901 the 1891 census, and for 1911 the 1901 census. Data is at the level of births, differentiated by household unit, year, and parish of birth. The dependent variable is the Working Class Score for the child, calculated using the data described above. All models include controls for sex and birth order, and fixed effects for the household unit, (1), (2), (4)–(6) and (8) include birth year fixed effects, (2) and (6) include parish of birth fixed effects, (3) and (7) parish of birth by year of birth fixed effects which subsume birth year and parish fixed effects, and (4) and (8) fixed effects for the parish of birth interacted with the father's parish of birth, which subsume parish fixed effects. Standard errors clustered by household and parish in parentheses. The number of observations in all models is 25,773,095. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.2: Within-father relationship between class distance and working class naming, calculating working class scores with 1901 or decade-specific data

	Working class score: County data				Cultural province data			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative class distance	-2.194^{**}	$-0.673^{\dagger}$	-0.985^{**}	1.239	-1.193^{**}	-0.925^{**}	-0.742^{**}	-1.027
	(0.403)	(0.405)	(0.328)	(1.257)	(0.340)	(0.352)	(0.263)	(1.169)
Fixed efffects: Household	x	x	x	x	x	x	x	x
- Year	x	x		x	x	x		x
- Parish		x				x		
- Parish × year			x				x	
- Parish × father's parish of birth				x				x
Controls: Sex, birth order	x	x	x	x	x	x	x	x
R^2	0.432	0.433	0.451	0.475	0.454	0.455	0.470	0.492

This table presents evidence of the relationship between distance to the working or upper class and working class naming, calculating working class scores using data from the 1881 census specific to the child's county of birth (1–4) or the groupings of counties classified as "cultural provinces" by the historian Phythian-Adams (5–8). Data is at the level of births, differentiated by household unit, year, and parish of birth. The dependent variable is the Working Class Score for the child, calculated using the data described above. All models include controls for sex and birth order, and fixed effects for the household unit, (1), (2), (4)–(6) and (8) include birth year fixed effects, (2) and (6) include parish of birth fixed effects, (3) and (7) parish of birth by year of birth fixed effects which subsume birth year and parish fixed effects, and (4) and (8) fixed effects for the parish of birth interacted with the father's parish of birth, which subsume parish fixed effects. Standard errors clustered by household and parish in parentheses. The number of observations in all models is 25,496,960. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.3: Within-father relationship between class distance and working class naming, calculating working class scores with region-specific data

	Irish name score				Working class score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative class distance	0.214	0.233	0.217	−0.850	−0.785**	−0.830**	−0.534*	−1.394*
	(0.318)	(0.353)	(0.330)	(1.102)	(0.219)	(0.234)	(0.223)	(0.583)
Excluding Irish names					x	x	x	x
Fixed efffects: Household	x	x	x	x	x	x	x	x
- Year	x	x		x	x	x		x
- Parish		x				x		
- Parish × year			x				x	
- Parish × father's parish of birth				x				x
Controls: Sex, birth order	x	x	x	x	x	x	x	x
R^2	0.416	0.417	0.431	0.453	0.482	0.482	0.494	0.516

This table presents evidence of the relationship between distance to the working or upper class and distinctively Irish and working class names. Data is at the level of the birth-by-household-by-parish-by-year. In (1)–(4) the dependent variable is the Irish name score of the child, calculated as for working class scores with 1881 census data, but identifying names over-represented among those born in Ireland. In (5)–(8), the dependent variable is the child's working class score, dropping names with Irish name scores of more than 80. All models include controls for sex and birth order, and fixed effects for the household unit, (1), (2), (4), (5), (6) and (8) add birth year fixed effects, (2) and (6) parish of birth fixed effects, (3) and (7) parish of birth by year of birth fixed effects which subsume birth year and parish fixed effects, and (4) and (8) add fixed effects for the parish of birth interacted with the father's parish of birth, which subsume parish fixed effects. Standard errors clustered by household and parish in parentheses. The number of observations is 25,773,095 in models (1)–(4) and 25,020,082 in models (5)–(8). $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.4: Within-father relationship between class distance and Irish naming, and working class identity excluding distinctively Irish names

	Working class score									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Relative class distance	−8.222**	−4.774*	−8.679**	−5.148**	−8.664**	−5.129**	−5.008**	−3.576**	−4.895**	−3.621**
	(2.107)	(2.044)	(0.775)	(0.619)	(0.814)	(0.651)	(0.846)	(0.805)	(1.105)	(1.026)
Fixed efffects: Year	x	x	x	x			x	x	x	x
- Father's occupation		x		x		x		x		x
- Parish			x	x					x	x
- Parish × year					x	x				
- Parish × father's parish of birth							x	x		
Controls: Sex, Birth order	x	x	x	x	x	x	x	x	x	x
Sample: Only non-movers									x	x
R^2	0.021	0.052	0.066	0.083	0.082	0.098	0.147	0.160	0.072	0.087

This table presents evidence of the relationship between distance to the working or upper class and working class naming. Data is at the level of births, differentiated by father, year, and parish of birth. The dependent variable is the Working Class Score for the child. All models control for sex and birth order. (1)–(4) and (7)–(10) include year of birth fixed effects, (3), (4), (9) and (10) include parish fixed effects, (5) and (6) parish-by-year fixed effects, which subsume parish and year fixed effects, (7) and (8) parish-by-father's parish fixed effects, which subsume parish fixed effects. Even-numbered models add fixed effects for the father's occupation. (9) and (10) are restricted to households where all children are born in the father's parish of birth. (7)–(10) can be considered generalized difference in differences, in which the unit is the father's parish-birth parish combination and the period is the year of birth. Standard errors clustered by household and parish in parentheses. The number of observations is 25,773,095 in models (1)–(8) and 8,665,426 in (9) and (10). $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.5: Relationship between class distance and working class naming, without household fixed effects

instead reflect aspirations for social mobility. Parents who observe more culturally similar individuals in the upper or middle class might infer that their children have greater chances of rising into that social group and therefore choose names associated with that milieu—either to facilitate or to signal such a transition. If this mechanism were driving our results, we would expect the effect of relative class distance on naming to be stronger among households already closer to the upper or middle class, for whom such behavior would be more plausible. Figure B.1 replicates our main specification by subsetting fathers into quintiles of occupational status, as measured by HISCAM. We find no significant differences in naming responses across fathers of different statuses. While not definitive, this pattern suggests a limited role for upward-mobility aspirations in explaining our results.[16] Similarly, in Table B.6, we re-estimate the models from Table 2, subsetting by the father's occupational class. We find on the whole slightly larger, but less precisely-estimated coefficients for upper class households, which make up a minority of the sample.

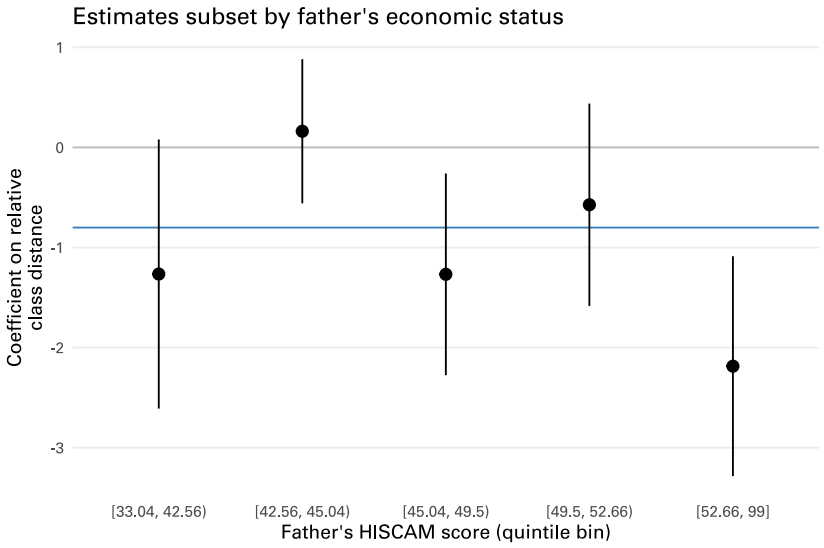


Figure B.1: Estimated relationship between relative class distance and working class identity, subset by economic status

Each point is the coefficient from a regression of working class scores against relative class distance with father and year fixed effects, subset to fathers with HISCAM scores in a given quintile of the distribution. Errorbars show 95% confidence intervals, constructed using standard errors clustered by father and parish. The blue line shows the full sample estimate.

We also consider the possibility that relative class distance may capture distance in occupation: people from culturally distant backgrounds might sort into different occupations. If the working class is more distant in terms of the occupations they hold (e.g., agricultural rather than industrial jobs), this occupational distance, rather than cultural distance, could dampen working-class identity. This channel is conceptually related to ours, as it also operates through distance, but along occupational rather than cultural lines. Table B.7 shows that relative class distance is not consistently correlated

[16]Another interpretation is that naming reflects imitation of names that are more common among the occupational class most culturally similar to the household. In other words, people may adopt names used by those of a similar cultural background, who also tend to belong to higher social strata. This interpretation is fully consistent with our theoretical framework: greater perceived similarity with an occupational class group fosters behavioral conformity, which is one manifestation of stronger identification with that group.

	Working class score							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative class distance	-0.629^{**}	-2.399^{**}	-0.793^{**}	$-1.769^{\dagger}$	-0.506^{*}	-1.119	-1.716^{**}	-0.268
	(0.232)	(0.882)	(0.249)	(0.952)	(0.237)	(1.072)	(0.621)	(2.535)
Father's occupational class:	Working	Upper	Working	Upper	Working	Upper	Working	Upper
Fixed efffects: Household	x	x	x	x	x	x	x	x
- Year	x	x	x	x			x	x
- Parish			x	x				
- Parish × year					x	x		
- Parish × father's parish of birth							x	x
Controls: Sex, Birth order	x	x	x	x	x	x	x	x
R^2	0.460	0.480	0.460	0.484	0.476	0.542	0.497	0.547

This table presents evidence of the relationship between distance from the working or upper class and working class naming. Data is at the level of births, differentiated by father, year, and parish of birth. Odd-numbered models are subset to households where the father's occupation is classified as working class by the Registrar-General, even-numbered models are subset to households where the father's occupation is upper or middle class. The dependent variable is the Working Class Score for the child. All models control for sex and birth order and include household unit fixed effects, (1)–(4) and (7)–(8) add birth year fixed effects, (3)–(4) include parish of birth fixed effects, (5)–(6) parish of birth by year of birth fixed effects which subsume birth year fixed effects, and (7)–(8) fixed effects for the parish of birth interacted with the father's parish of birth, which subsume parish fixed effects. Standard errors clustered by household and parish in parentheses. The number of observations is 20,255,111 in models subset to working class households and 1,748,941 in models subset to upper class households. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.6: Within-household relationship between class distance and working class naming, subset by the father's occupational class

	Share with same occupation				Working class score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative class distance	−0.008	−0.055**	−0.060**	−0.004	−0.533*	−0.674*	−0.299	−2.085**
	(0.008)	(0.004)	(0.003)	(0.013)	(0.263)	(0.285)	(0.269)	(0.745)
Share with same occupation					0.527**	0.253*	0.355**	0.083
					(0.094)	(0.114)	(0.125)	(0.154)
Fixed efffects: Household	x	x	x	x	x	x	x	x
- Year	x	x		x	x	x		x
- Parish		x				x		
- Parish × year			x				x	
- Parish × father's parish of birth				x				x
Controls: Sex, birth order	x	x	x	x	x	x	x	x
R^2	0.939	0.957	0.968	0.976	0.469	0.469	0.484	0.500

This table presents evidence of the relationship between distance from the working or upper class and the share of working class adult men (age ≥ 16) in the parish with the same occupation as the father. Data is at the level of the birth-by-household-by-parish-by-year. In (1)–(4) the dependent variable is the share of working class men in the parish at the closest census with the same occupation as the father, and throughout observations are subset to families in which the father is employed in a working class occupation. In (5)–(8), the dependent variable is the child's working class score, with the share of working class men in the same occupation as the father included as a covariate. All models include controls for sex and birth order, and fixed effects for the household unit, (1), (2), (4), (5), (6) and (8) add birth year fixed effects, (2) and (6) parish of birth fixed effects, (3) and (7) parish of birth by year of birth fixed effects which subsume birth year and parish fixed effects, and (4) and (8) add fixed effects for the parish of birth interacted with the father's parish of birth, which subsume parish fixed effects. Standard errors clustered by household and parish in parentheses. The number of observations in all models is 18,863,500. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.7: Within-father relationship between class distance and share of working men in parish with same occupation

with the share of working-class adults in the same occupation (columns 1–4). When we control for the share of working-class adults in the same occupation, that measure itself increases identification with the working class. Relative class distance, however, remains predictive, suggesting that both forms of distance matter for class identification.

Might the decrease in working class identity we identify also reflect stronger national identification? In Shayo (2009), higher perceived distance from the working class is associated with greater identification with the nation. Although this mechanism may operate in our context as well, our emphasis is on the choice between upper- and working-class identity. National identification would only pose a problem for our interpretation if it were confounded with upper-class identity. Fouka and Serlin (2024) show that names distinctive of London and the Southeast—regions historically richer and with names likely overrepresented in upper class occupations—spread across England and Wales during the Second Industrial Revolution, supplanting local naming patterns and identities.

To examine this concern more systematically, we construct a measure of local identity analogous to our working class scores, capturing the relative frequency of names in each county compared to the overall population. Columns 1–3 of Table B.8 show that higher relative class distance lowers

	County score: father's	child's county		Working class score		
	(1)	(2)	(3)	(4)	(5)	(6)
Relative class distance	-0.812^{**}	-3.104^{**}	-3.824^{**}			
	(0.267)	(0.499)	(0.326)			
County score (father's county)				0.025		
				(0.018)		
County score (child's county)					0.001	0.001
					(0.024)	(0.025)
Fixed efffects: Hhd, year	x	x	x	x	x	x
- Parish			x			x
Controls: Sex, birth order	x	x	x	x	x	x
N	25761006	25761006	25761006	31960005	32082516	32082516
R^2	0.444	0.441	0.446	0.456	0.456	0.457

This table presents evidence of the relationship between distance to the working or upper class and local identity. Data is at the level of births, differentiated by household, year, and parish of birth. In model (1) the dependent variable is the child's county score, calculated analogously to working class scores but using the relative frequency of names among those born in the father's county of birth as a measure of identity. In (2)–(3), the equivalent but for the child rather than the father's county of birth. In (4)–(6) the dependent variable is the child's working class score. All models include fixed effects for household and birth year and controls for sex and birth order. (3) and (6) also include fixed effects for the parish of birth Standard errors clustered by father and parish in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.8: Within-father relationship between class distance and local identity

local name scores, suggesting that when people are more culturally distant from the working class, they identify less with local culture. This pattern is consistent with greater identification with national, rather than local, identities, as in Shayo (2009). However, this channel does not explain our main results: columns 4–6 of Table B.8 show that working class scores are uncorrelated with local identity. In other words, the class valence of names does not simply capture their local versus national content. When we flexibly control for local identity in Table B.9, our estimates remain unchanged.

	Working class score				
	(1)	(2)	(3)	(4)	(5)
Relative class distance	-0.792^{**}	-0.784^{**}	-0.812^{**}	$-0.408^{\dagger}$	-0.653^{**}
	(0.223)	(0.244)	(0.208)	(0.212)	(0.209)
County score (father's county)	0.030				
	(0.019)				
County score (child's county)		0.008			
		(0.025)			
Fixed efffects: Household,					
year, parish	x	x	x	x	x
Controls: Sex, birth order	x	x	x	x	x
- County score × father's county			x		x
- County score × child's county				x	x
N	25761006	25761006	25761006	25761006	25761006
R^2	0.471	0.470	0.538	0.545	0.553

This table presents evidence of the relationship between distance from the working or upper class and working class identity, controlling for local identity. Data is at the level of births, differentiated by household, year, and parish of birth. The dependent variable is the child's working class score, the independent variable relative class distance. (1) controls for the child's county score for the father's county of birth, (2) for the child's county score for the child's county of birth, (3) and (5) control for the child's county score for the father's county, allowing the slope to vary by the father's county, (4) and (5) for the county score for the child's county, allowing the slope to vary by that county. All models include fixed effects for household, parish and birth year, and controls for the child's sex and birth order. Standard errors clustered by household and parish in parentheses. $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table B.9: Robustness of estimates to controlling for local identity

C validating upper class rhetoric

C.1 *Election address with least upper class rhetoric*

HAMPSTEAD ELECTION, 1895. TO THE ELECTORS OF THE BOROUGH OF HAMPSTEAD. GENTLEMEN, HAVING been invited by the Conservative and Liberal Unionist Associations to stand as their candidate at the approaching General Election. I once more ask you to give me a renewal of your confidence and return me as your member to the House of Commons. The late Government have for three years pursued a policy of attack on the most cherished institutions of the country. They have assailed, or threatened, the integrity of the United Kingdom, the Church, the rights of property, and the House of Lords. Their Foreign policy has been marked by weakness and vacillation, and under their rule the country has been overshadowed by depression and anxiety. I am not in a position to formulate in detail the policy which will be pursued by the present Government should they (as we all expect) find themselves at the head of a substantial majority in the new Parliament, but the general lines of that policy have been clearly indicated. They will steadfastly resist Home Rule, which is still retained as the foremost item in the programme of the Radical Irish Coalition, but we may expect from them in due time a measure of Local Government for Ireland. The Church will be safe from attack, and will be able to pursue its beneficent spiritual and educational work without hindrance. They will promote legislation having for its object the well-being of the people. They will administer the affairs of the nation with Efficiency and Economy. Under the auspices of Lord Salisbury we may confidently rely on the interests and honour of the country being safeguarded abroad; while the appointment of the Duke of Devonshire to be the head of a Council of National Defence is a guarantee that the strength of England will be efficiently organised. It must be a source of satisfaction to all Unionists to see that the friendly alliance between Conservatives and Liberal Unionists, which has lasted for nine years, is to be further cemented by the inclusion of the most prominent men from the make of the two parties in the new Government, which promises to be one of the strongest of the century. I appeal to you with confidence to do your share towards giving the Government the necessary support in Parliament, by returning me as your member. Should you do so, it will always be my pleasure, so it is my duty, to do all in my power to further the interests of the Borough and to be useful to my constituents, without distinction of party. With grateful thanks for the kindness and consideration shown to me in the past, I remain, Your obedient servant, E. BRODIE HOARE. 109, St. George's Square, S.W. July, 1895.

C.2 *Election address with 5th percentile upper class rhetoric*

TO THE ELECTORS OF THE STRAND. GENTLEMEN,- Parliament having been dissolved, I have once more the privilege of asking you to accord me your support and confidence at the coming Election. The Conservative and Unionist Government which recently retired from Office, was especially charged with a mandate to conclude the South African War, and to reorganise the territories which the results of that war had added to the Empire. Already much progress has been made in the direction of Self- Government in the case of the Transvaal, and I am in favour of granting both to that Colony and to the Orange River Colony the fullest forms of Responsible Self-Government through representative Institutions as soon as such privileges can be wisely granted in the interests of the Empire and of the Colonies themselves. The introduction of Chinese Coolies into the Transvaal has met with fierce opposition and misrepresentation in this country. It does not appear that this agitation has been in any way justified by the results of the experiment, and I am quite content to leave the inhabitants of the Transvaal to form their own opinion and pass their own verdict upon it. I believe that the Colonial Policy of the late Government has done much to

encourage closer relations between the Mother Country and the Colonies, and that both at home and abroad the British Empire is acknowledged to be bound together by ties which have withstood the strain of War and which have been drawn closer by a clearer appreciation of the necessity for Imperial Unity. The system of Colonial Conferences instituted in 1887 by a Conservative Government and continued in 1897 and 1902 provides the means for periodical exchanges of views on important Imperial matters between the Home and Colonial Governments. I am in favour of placing such conferences on a more systematic footing as suggested in a recent dispatch by the late Colonial Secretary. I am also prepared to support the proposal that the Trade Relations between the Colonies and the Home Country shall be the subject of a special and unfettered Conference, for, although I am not one of those who think that any system involving the protective taxation of any kind of raw material will in the end improve our Imperial Trade, I still consider that to refuse our Colonies any discussion upon the question would be unreasonable, more especially as they are aware that opinion at home is not unanimous. I consider that the conclusions of such a Conference, if involving a change in our Fiscal System should be submitted to the test of a General Election before being acted upon. Attention has also been drawn to the fact that in negotiating Commercial Treaties with Foreign Governments our own Government being bound by the Fiscal Policy which has been in existence for many years is unable to negotiate with a free hand; the policy of retaliation is proposed as a remedy for this difficulty. In an address to the Electors of the Strand it is not possible to discuss the methods by which such a policy may be carried out, but, provided that such retaliatory duties, if proposed, are in each case submitted to Parliament for approval, I shall be ready to consider each one on its merits, with due regard to the interests of our commerce as a whole. During the past five years, important changes have been made in the distribution and organisation of our Fleets at home and abroad; while resulting in a large diminution of the Naval Estimates to the extent of several millions annually, these changes have also increased the efficiency of the service which it must always be the object of the Unionist Party to preserve at the highest point. In the Army, also, alterations have been made, and it cannot be denied that the Private Soldier is now better housed, fed, and paid than ever before. The creation of a General Staff of the Army, which has not hitherto existed, has been welcomed by expert opinion as being calculated to improve the organisation of the Service and to increase its efficiency for war. I am in favour of increasing the efficiency of the volunteer forces and of giving them every inducement to render themselves fit for active service when needed. The Foreign Policy of the late Government has been proved eminently successful; very cordial relations have been established with France and Japan, and a Treaty has been signed, which will have a most important influence in preserving Peace in the Far East for many years to come. In Domestic Legislation, two of the most important Acts passed have been the Licensing Act and the Education Act I object to any attempt to amend the principles of either of these measures. As in the past, I am now opposed to the large schemes of Municipal Trading put forward by the London County Council. I welcome the appointment of a Commission to enquire into the Poor Law and the causes of Unemployment, and, as I have stated in former Addresses, I am in favour of more careful classification of Workhouse Inmates. I should be glad to see swept away many of the anomalies which now exist in Electoral Areas, and I consider that a Redistribution Bill is urgently needed for this purpose. Gentlemen,- If you are pleased to return me once more as your Representative. I shall cordially support our Leader as a member of the Unionist Party. I have no doubt but that Home Rule for Ireland must force itself upon the present Radical Government, if they are confirmed in power, and recent events have shown that, either slowly or quickly, it is the policy of the present Prime Minister to grant the demands of Irish Separatists. To this policy I shall offer my strenuous opposition. The special work of the Unionist Party in the Strand is therefore not concluded, and I trust that for another Parliament you will permit me to serve your interests to the best of my ability.

I have the honour to remain, Gentlemen, Your obedient Servant, W. F. D. SMITH. 186. Strand.

C.3 *Election address with 10th percentile upper class rhetoric*

TO THE ELECTORS OF THE HARROW DIVISION OF MIDDLESEX. GENTLEMEN. On the invitation of Liberal and Labour Organisations in the Constituency, I have the honour to offer myself to represent your important Division in the House of Commons. The occasion is a critical one. Freed from the anxiety (though not from the burden) of the South African War, men's minds revert to questions touching the welfare of our people. Great improvement has been wrought in our material prosperity, largely owing to Free Trade, in the last half- century; but when Free Trade has done its best, it has not solved our Social problems, and alone it cannot solve them. What shall be the next step? The Conservatives say 'backwards'; I say, 'forward.' Their leaders have put the reversal of our Free Trade practice in the forefront of their policy, and have thus made our Fiscal policy the dominant factor at this election. It is too large to be relegated to a second place. Though Free Trade is not all we need, it is nevertheless essential to the well- being of the people. Its tendency has been (in accordance with the anticipations of half- a- century ago) to raise wages, cheapen the necessaries and comforts of life, shorten the hours of labour, increase wealth, and generally promote the well- being of the nation. Free Trade. In place of Free Trade, Conservatives offer us preference, retaliation, or undisguised protection. To expect from these anything better than diminished trade, harder conditions of life, friction with, the Colonies, fiscal wars with other countries, and political corruption at home, is to ignore history and close our minds to reason. I hold firmly by Free Trade, and rejoice that we as a nation have the honour of having so long held aloft the light of a truth which shall yet prove to be for the healing of the nations. Still we have poverty and unemployment- too much poverty; too much; because partly preventable. If it were due to a stinginess of nature, we should have to bear it or thin our numbers by emigration. But while our land laws are what they are; while in education we are surpassed by some other countries; and while our licensing laws are conceived in the interests of a class rather than in those of the community, we have not done our best to eradicate poverty. Land. The estrangement of the people from the land is a serious feature of modern life, resulting in city slums, and in physical, mental and moral deterioration. It is worth while to make a sacrifice to restore the people to the land, and yet perhaps no great sacrifice is needed, only common sense and justice. It is not common sense that land should be under the control of dead men, and ??? maintenance of ??? families but unfavourable to the general weal. Land should be for the unfettered use of the living. Land values are a proper subject for taxation. Rent is individual wealth, but not national wealth. It is a deduction from the fruit of a man's labour for the benefit of the landlord. This tendency has been kept in check in agricultural districts by the opening of our ports. Had it not been for the abolition of the Corn Laws, we should probably before this have had a starving population or a revolution. In our growing towns the tendency to higher rents is very marked. Some part of the increase should, by the taxation of site values, be secured for the community, which by its presence creates the value. By the taxation of land, whether it be in use or idle, we should relieve the burdens of the people, bring more building land into the market, provide better housing accommodation, and have something over with which to begin the desirable work of securing the land for the nation. I believe that until we amend our land laws we are handicapped in our fight against poverty. Education. What other nations can do we can do, and we should not be behind any in the efficiency of our education. The late Acts cannot give us efficiency. They are too unjust to command the co- operation of all citizens. The people who pay should control, and the more directly the better, and their choice of head teachers should not be limited by law to the members of any particular religious denomination. This does not mean the exclusion of religious teaching from the schools. For thirty years we have had taught, to the apparent

satisfaction of parents, a body of religious and moral truth, not professing to be exhaustive, yet feasible and valuable, and capable of falling into place with other more distinctive instruction given in Sunday Schools and Churches. I should regard the exclusion of the Bible from our schools as a national calamity. Licensing. The Act of 1904 was not just to the nation. I would amend it by imposing a time limit to the operation of the Compensation clauses. The people should be consulted in their districts as to the number of licensed houses they require. I am in favour of closing public houses on Sundays and election days. National Expenditure. The enormous increase in our national expenditure, equal to about 50 per cent. in ten years, is alarming: and it will be the duty of the new Government and the new Parliament to effect such economies as are consistent with national needs. Sound finance is essential to national prosperity. Trade Unions It would make the law what the Act of 1871 intended, but failed, to make it. For thirty years the Act, as it was wrongly understood, worked well; and now that the law has been declared in the Taff Vale case to be different from what was intended, the mistake should be rectified. Ireland. I hold by two governing principles:- (1) The British Parliament must be paramount and the Union must be maintained and safeguarded; and (2) To safeguard the Union we must respect Irish sentiments, and let Irish ideas control the management of purely Irish affairs. The Union is safer when its value is recognised by the Irish themselves. There is no reason in the nature of things (apart from grievances) why Ireland should be more desirous of separating from Great Britain than is Great Britain of separating from Ireland. The Colonies. I rejoice in the strength of the bond which unites us, and I am convenced that it is in a large measure due to their rights of self- government, and to their and our fiscal freedom. As regards the Transvaal and the Orange River Colony, I look forward to an early application of the talisman of self- government, to the creation of a parliament elected on a wide suffrage, to deal with Chinese Labour and all other local questions. Foreign Policy. If we can cultivate and exhibit an unaffected good- will to all peoples and be sympathetic with the oppressed, we shall not lack friends, and we shall, I hope, be fostering the increasing love of peace and horror of war which are discernible in the civilised world. Entangling alliances I dislike, If they must be, let them be very definite and of short duration. Of the principle of arbitration I cordially approve. To summarise: I am in favour of Liberty and Equality, civil and religious; of Electoral Reform; of Women's Suffrage; of Land, Licensing and Educational Reform; of Free Trade; of Economy; of Peace; of Justice to Ireland; of Cordial Relations with the Colonies; and of all sound measures for the improvement of the condition of our people which will not produce greater evils than they cure. One hope I venture strongly to express, that in this most critical election there will be but few abstentions from the poll. I have the honour to be, Gentlemen, Your faithful servant, JAMES GIBB. 52, Willesden Lane, Kilburn, N.W. January, 1906.

C.4 *Election address with 90th percentile upper class rhetoric*

TO THE ELECTORS OF THE CLITHEROE DIVISION. GENTLEMEN, The unanimous invitation of the Unionist Three Hundred, as well as the hearty reception I have met with on all sides, emboldens me to ask for the favour of your support. I am a Lancashire man familiar with the feelings and requirements of Lancashire men. If I am sent by you to Parliament I will do all in my power to merit the high distinction conferred upon me. The aim of a Representative of the people should be to secure the greatest happiness for the greatest number; to avoid 'the falsehood of extremes' and to take care that in the onward march of National Progress the interests and liberties of individuals and classes are not unheard, ignored, or ruthlessly trampled down. By means of a Committee Elected by all Ratepayers I would place the Veto or Granting of Liquor Licenses in the Hands of the People. The Unionist Government have helped countless thousands of toilers by the great boon of Free Education. I object to taking away with one hand that which has been given by the other, and I

should strenuously oppose all attempts to retard the time at which a child may commence to earn his own living and become a help to hard- worked and often overburdened parents. I consider EIGHT hours affair day's work for all engaged in mining industries, and I would assist any proposal to ensure the same; and as regards the textile and other trades I should be guided by the feeling of the workers themselves. The Laws relating to the land require Reform Transfer should be simplified: Registration rendered easy and obligatory, and the questions of Mining Royalties and Copyhold Enfranchisement dealt with. Economical administration of Finance is the attribute of every well- governed state; but Our Island Home, our Colonies, and the Commerce of our vast Empire must be rendered safe from attack, and I should not grudge money well and carefully spent upon the Naval and Military forces of our Country. I am happy to think that under a Unionist Government taxation has been reduced, and Local Burdens relieved by grants in aid, while the Local Government for England Act, 1888, has given you the full control and administration of your local affairs. Gentlemen,- - I am prepared to extend a similar measure to Ireland, but believing as I do that the Imperial Parliament has shown itself both able and willing to deal with all Irish questions, I am totally opposed to the creation of an Irish Parliament nominally subordinate to but practically independent of the Imperial Parliament at Westminster. Such a measure would, I fear, be fertile in Civil and Religious Feuds, and fraught with disaster to the Industrial undertakings of that country. Gentlemen,- - You have, on grounds of public policy, abolished one religious ascendency in Ireland, are you now prepared to establish another? Will you surrender Ireland to the Government of Faction in place of Government by the Queen? For my part I say 'Never!' I am, Gentlemen, Yours faithfully, W. E. BRIGGS.

C.5 *Election address with 95th percentile upper class rhetoric*

TO THE ELECTORS OF THE PARLIAMENTARY BOROUGH OF MERTHYR TYDFIL. GENTLEMEN,- A new Parliament is about to be elected, and I respectfully ask you for a renewal of the confidence you have so generously extended to me during the past twelve years, and to again do me the honour of returning me to represent in the House of Commons the constituency in which I was born and bred. The Dissolution of Parliament at such a time on so old a register is without precedent in the political history of the constituency, and will involve the virtual disfranchisement of between three and four thousand electors in Merthyr and hundreds of thousands throughout the country. No serious attempt has been made by the Government or its supporters to justify or explain this extraordinary proceeding, and we are forced to the conclusion that the Tory party fear the verdict of a calm and considered judgment of public opinion upon a full register, and hope to snatch victory by exploiting for party ends the national outburst of patriotic enthusiasm that has been aroused by the pluck and endurance of our gallant troops in South Africa. The war is practically over, but the successes of our brave fellow countrymen have been won in spite of the blunders of the War Office. The Government can claim no share of the credit. This war was rendered inevitable by the bungling diplomacy of the Colonial Office, but once it had broken out it was necessary to go through with it to a successful issue, and I regard the annexation of the two Republics as a consequence that could not be avoided. So little foresight did the Government exhibit that they thought the war would be over in a few weeks, and that twenty thousand troops would be sufficient. It has taken ten times that number nearly twelve months to bring it to a conclusion at the expense of forty thousand men killed, wounded, and invalided, and a cost in money that will not fall far short of one hundred million pounds. The Hospital arrangements have constituted a scandal, and far more of our gallant fellow countrymen have succumbed to disease than to wounds. We have been told that the Tory Government has a right to determine the issue upon which the election shall be fought, and that that issue is the policy that dictated the war and rendered it inevitable. I can well understand Mr Chamberlain's anxiety to get himself whitewashed and to cover over the failure to

redeem the pledges made in 1895. But I deny the right of the Tories to dictate to the electors the issue upon which they shall vote, and I for my part decline to fight upon an issue that is dead, a war that is ended, or annexation that has already taken place. The verdict of the electors cannot revive the thousands of lives that have been sacrificed or restore the millions of public money squandered. The Government have thrown out the challenge; the right of selecting the weapon rests with the challenged. I propose taking the whole record of the Tory party- a record of large promises and small performances, of bluster and timidity, of spendthrift finance, of reaction at home and loss of prestige abroad. I hold, gentlemen, to my old programme, the maintenance and consolidation of our great Empire, and upon my banner you will find inscribed Home Rule, Religious Equality, Temperance Reform, Old Age Pensions, Land Law Reform, Taxation of Ground Bents and Royalties, Abolition of the System of Tied Houses, One Man One Vote and Every Man a Vote, the Extension of the Franchise to Duly Qualified Women, and the Abolition of the Hereditary Principle in the Upper Chamber. I shall, if returned, support an eight hours day for miners from bank to bank, a extension and amendment of the Compensation. Act in the direction of payment for the first fort- night of incapacity from accident, and the allowing of compensation in all cases of accident. It has always been my object to safeguard and promote the interests of the working classes, and if you honour me with your confidence I shall continue to do so, and will at all times give special attention to questions affecting Wales and to all matters pertaining to the staple industries of the district. I am, Gentlemen, Your obedient servant, D. A. THOMAS. Yecyborwen, September 24th, 1900. 9102

C.6 *Election address with most upper class rhetoric*

GENERAL ELECTION, 1900. TO THE ELECTORS OF OWM RHONDDA. DEAR FRIENDS & FELLOW WORKMEN, It makes it imperative on me once more to appeal to you for your generous sympathy as you have always tendered my past 15 years' unbroken representation. The imperativeness arises from the following:- 1st. Because of the Government having forced upon the country an unprecedented premature dissolution, which will disfranchise about half a million British workmen. 2nd. Because of the Rhondda Conservative Club Association having forced upon us an unnecessary, hopeless contest, which will cost hundreds of pounds. 3rd. Because of myself having been again invited to fight the Labour and Liberal battle in the constituency. 4th. Because I have accepted the invitation. I have no new Programme that would be necessary. The Rhondda Programme has always been a comprehensive one. It comprises Labour and Liberalism to the fullest extent; and in two words these they are- 1st. The Rights of Labour. 2nd. The Will of the People. Yours obediently, MABON.

D DECOMPOSING HETEROGENEITY IN CLASS DISTANCE CHANGES

Figure 7 shows that over the 1891–1911 period, places with low agricultural employment experienced increases in relative distance from the working class, while in places with high agricultural employment, relative distance from the working class increased. Figure A.5 shows that this change in the spatial distribution of relative class distance is due to rising distance from the working class in initially agricultural areas, and rising proximity to the working class in initially less agricultural areas. This appendix further decomposes this change in the spatial distribution of class distance.

We see similar patterns looking at the share of working class residents born outside the county, a cruder measure of cultural proximity, in the left panel of Figure D.1. In 1891, places with a large share of employment in agriculture have more homogeneous working classes, drawing a larger share from the same county. Between 1891 and 1911, this relationship flattens. The share born in county increases in places with low initial agriculture and decreases in places with high initial agriculture.

Places with low to intermediate agriculture experience population growth, while population in the most agricultural regions is static (Figure D.1, right panel).

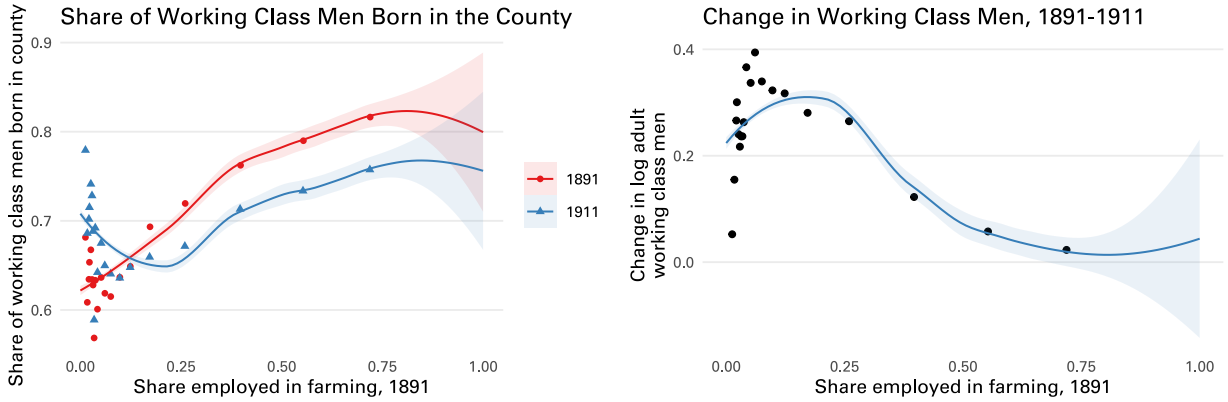


Figure D.1: In 1891, agricultural areas have the highest share of same-county residents, by 1911 this slope levels off, while agricultural areas experience the slowest population growth

The left panel plots the share of working class adult men who were born in the same county as the parish, in 1891 and 1911, against the share employed in farming in 1891. The right panel plots the change in the log number of working class men between 1891 and 1911 against 1891 farming. Dots are binned averages, lines are loess curves with 95% confidence intervals, all weighted by 1891 population.

What is this picking up? A falling share of residents born in the county could reflect two phenomena. First, it could be that the share of people born in the area who decide to stay falls. Second, it could be that the area grows and draws in more migrants from distant locations. Formally, we can decompose the share of the resident population born in the county into the ratio of residents born in the county relative to people in all locations born in the parish, a measure of how far the parish retains people born locally, and the ratio of current residents to people born in the parish, a measure of net migration:

$$\underbrace{\frac{\text{Residents born in county}_{it}}{\text{Residents}_{it}}}_{\text{Same county share}} = \underbrace{\frac{\text{Residents born in county}_{it}}{\text{People born in parish}_{it}}}_{\text{Population retention}} \Big/ \underbrace{\frac{\text{Residents}_{it}}{\text{People born in parish}_{it}}}_{\text{Net migration}}.$$

If a location experienced high rates of out-migration, we would expect the Population retention component to decrease, whereas if a location experienced population growth, bringing in migrants from distant origins, we would expect the Net migration component to increase. Taking logarithms and differencing between 1891 and 1911, we have

$$\Delta \log(\text{same county share}_i) = \Delta \log(\text{Population retention}_i) - \Delta \log(\text{Net migration}_i).$$

In words, the change in the log share born in the county can be expressed as the change in the rate at which locally born residents stay in the parish, minus the change in the net rate of migration into the parish. Figure D.2 plots these components against the share employed in farming in 1891, for the working class. In the areas with the least employment in farming, the increase in the share born in the county is driven by falling net migration (the red line), even while the share born in the area remaining in the parish declines. This change reflects growth in these areas levelling off. In 1881 these areas were drawing in many migrants from other locations, by 1911 the children of these

migrants, born in the area, made up an increasing share of the population. In the areas with the most employment in farming, relative out-migration dominates stable net migration, leading to a fall in the share born in the county, while at intermediate levels in-migration accounts for a fall in the share born in the county.

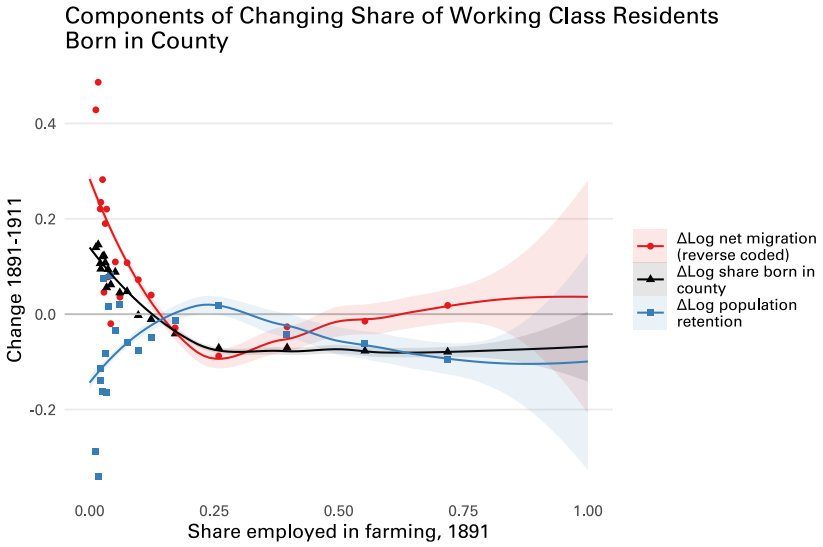


Figure D.2: In agricultural regions, out-migration drives rising working class distance, in urban areas falling net migration decreases working class distance

This figure plots the change in the log share of working class residents born in the county between 1891 and 1911, and the changes in the number of county born working class residents divided by the number born in the parish ("population retention") and the number born in the parish divided by the number of residents ("net migration"), which collectively sum up to the changes in the log share of residents born in the county. Dots are binned averages, lines are loess curves with 95% confidence intervals, all weighted by 1891 population.

Why were people migrating out of agricultural regions? Among the universe of working class men recorded as being born in a given parish in a given census, we can apply the law of total probability to write the probability of an individual staying within the county as the sum over occupations of the within-occupation rates of staying, multiplied by probability of taking each occupation:

$$P(\text{stay in county})_{it} = \sum_j P(\text{occupation} = j)_{it} P(\text{stay in county}|\text{occupation} = j)_{it}.$$

With some rearranging, we can then write the difference between 1891 and 1911 in terms of reallocation across occupations and within-occupation changes in the rate of people remaining in

the county:

$$\underbrace{P(\text{stay in county})_{i,11} - P(\text{stay in county})_{i,91}}_{\text{Total change}} =$$

$$\sum_j \underbrace{(P(\text{occupation} = j)_{i,11} - P(\text{occupation} = j)_{i,91})\, P(\text{stay in county}|\text{occupation} = j)_{i,91}}_{\text{Across-occupation changes}} +$$

$$\sum_j \underbrace{P(\text{occupation} = j)_{i,11}\, (P(\text{stay in county}|\text{occupation} = j)_{i,11} - P(\text{stay in county}|\text{occupation} = j)_{i,91})}_{\text{Within-occupation changes}}.$$

Expressing the change in terms of within-occupation or between-occupation changes helps us diagnose whether structural change in the economy accounts for these spatial patterns. As the share of employment in agriculture declines, we would expect people born in rural areas to shift out of agriculture and into occupations like manufacturing that were less concentrated in rural areas.

Figure D.3 plots all three components against 1891 agriculture. For areas with the least employment in agriculture, the reduction in population retention occurs entirely within-occupation. As the share in agriculture increases, there is a progressively larger reallocation across occupations, away from those in which people were more likely to stay within the county.

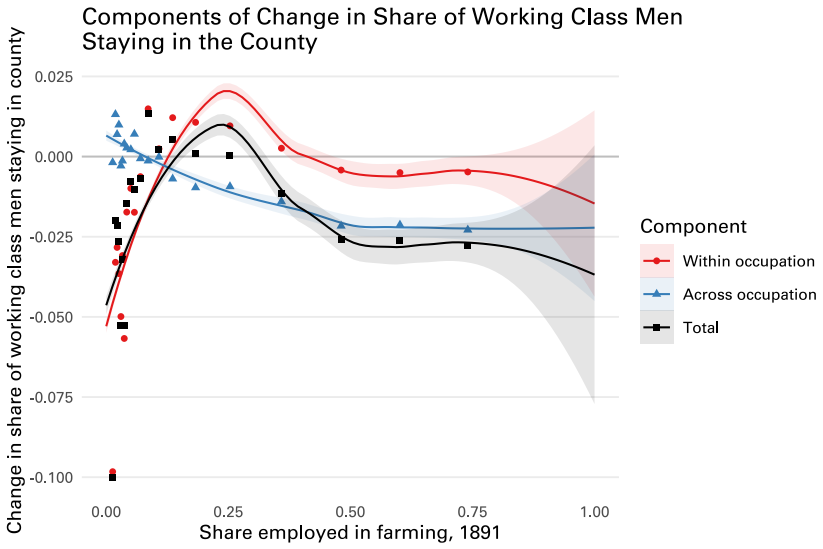


Figure D.3: In agricultural regions, falling population retention is attributable to shifts away from occupations with high retention rates, while in the least agricultural regions, the change is entirely within-occupation

This figure plots the change in the log share of working class residents born in the county between 1891 and 1911, and the changes in the number of county born working class residents divided by the number born in the parish ("population retention") and the number born in the parish divided by the number of residents ("net migration"), which collectively sum up to the changes in the log share of residents born in the county. Dots are binned averages, lines are loess curves with 95% confidence intervals, all weighted by 1891 population.

What are these occupations associated with lower or higher rates of out-migration? Among those born in agriculture-intensive areas, those employed in farming were less likely to migrate to other counties, relative to others born in the same parish. Figure D.4 plots the probability of a

working class man living in the county of birth, conditional on being born in a given parish and working in a given sector of the economy, against the agricultural share of the birth parish. As the national share of employment declines, people in agricultural regions face rising incentives to switch into other sectors, which often requires moving across regional boundaries.

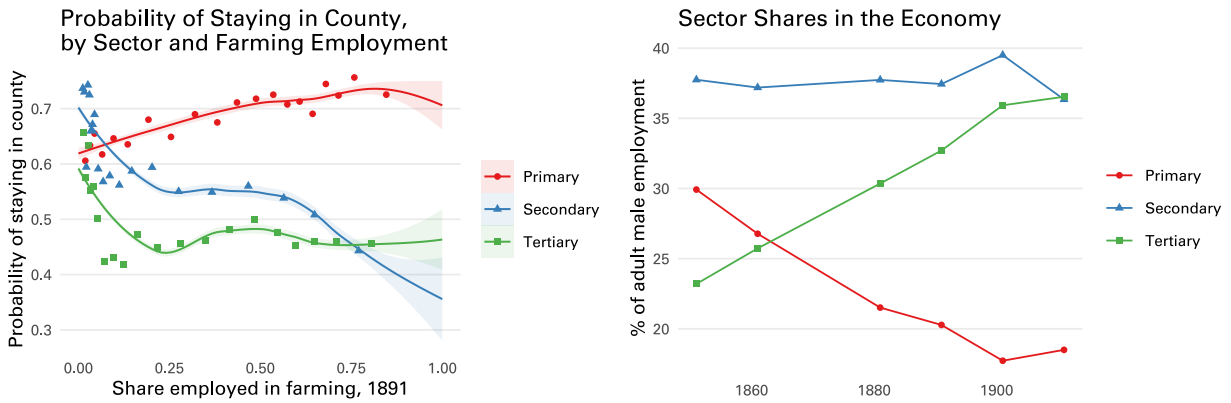


Figure D.4: In agricultural regions, people were most likely to remain in the region if they had agricultural jobs, these systematically declined

The left figure plots the share of working class men staying within their county of birth in 1891, by primary/secondary/tertiary sector against the share of farming employment in 1891 in their parish of birth. Dots are binned averages, lines are loess curves with 95% confidence intervals, all weighted by the number of observations born in each parish/occupational group combination. The right panel plots the share of adult men employed in each sector of the economy in censuses, 1851–1911.