

ARTICLE

Geography, ethnicity, genealogy and inter-generational social inequality in Great Britain

Paul Longley¹  | Tian Lan^{1,2}  | Justin van Dijk¹

¹Department of Geography, University College London, London, UK

²School of Resource and Environmental Science, Wuhan University, Wuhan, China

Correspondence

Tian Lan, School of Resource and Environmental Science, Wuhan University, 129 Luoyu Road, Wuhan, 430079, China.
Email: tian_lan@whu.edu.cn

Funding information

The research for this paper was supported by the Consumer Data Research Centre, which is funded by the Economic and Social Research Council (ES/L011840/1)

Abstract

This paper documents population-wide inequalities of outcome in Great Britain amongst and between long-established and more recently arrived family groups. ‘Establishment’ is defined using family group presence in the 1851 Census of Population as a benchmark, and the ethnicity or nationality of more recent migrants is determined through classification of given and family names. Inequalities of outcome are measured using a harmonised indicator of neighbourhood deprivation (hardship). White British individuals tend to live in the best neighbourhoods, but within-group inequalities reflect regional locations in which different family names were first coined 700 or more years ago. The living circumstances of White Irish and Chinese migrants are observed to be in line with long-established White British family lines, but other conventionally defined ethnic groups fare worse, some very markedly so. Disaggregation of conventional ethnic groups used by the Office for National Statistics such as White Other and Other Asian reveals stark within-group inequalities. These findings suggest: (a) regional origins of inter-generational inequalities amongst the White British; (b) comparable neighbourhood environments experienced by the White Irish, Chinese and some White Other groups and (c) significantly worse neighbourhood circumstances within and between other more recently arrived immigrant groups. This work has several implications for understanding economic assimilation of migrants and the existence of inequalities amongst and between populations.

KEYWORDS

deprivation, ethnicity, inequality, regional geography, social mobility

1 | INTRODUCTION

Increasing inequalities in wealth, income and access to life chances pose what may prove to be the most formidable challenges of the 21st century. Not only do they threaten the stability and cohesion of contemporary societies; they are also closely interlinked with other global problems, including housing market failures, the future of work and the energy

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

The information, practices and views in this article are those of the author(s) and do not necessarily reflect the opinion of the Royal Geographical Society (with IBG).

© 2023 The Authors. *Transactions of the Institute of British Geographers* published by John Wiley & Sons Ltd on behalf of Royal Geographical Society (with The Institute of British Geographers).

efficiency of 'net zero' cities. Perceptions of inequalities are often fundamentally geographical: for example, just 31% of people in England's north-east perceived good opportunities for social progression, compared with 74% of those in the south-east and 78% of Londoners (Social Mobility Commission, 2020a). Evidence of social equity inertia arises despite increasingly favourable education outcomes, funding of training and welfare support (Social Mobility Commission, 2019, 2020b). Recent academic work provides pointers to the selective accumulation, concentration and inter-generational transmission of financial and physical assets as a core driver of contemporary patterns of inequalities (Almeida, 2021; Clark & Cummins, 2014; Piketty, 2014). Housing, real estate and the financialisation of urban assets play particularly strong roles in the divergence of life chances (Harvey, 2010) and were further amplified through expansive monetary policies and slow economic growth in many countries after the 2009 financial crisis. A further wave of asset inflation and labour market polarisation was unleashed by the 2020 Covid pandemic, accentuated by differentials in the stability of employment and socially selective flexibility of homeworking. Together these outcomes affect the efficiency with which the most desirable and sustainable urban space is occupied, manifest in the contrast between increasing inter-generational crowding of some private rental space and relative under-occupation of late life-stage owner occupied stock. Further residential disruption arises from short-term letting (Airbnb) and absent owners of residential investments. Policy inertia also has important implications for urban energy transitions, which are reinforced by perverse property taxation regimes.

These issues and requirements have gathered urgency over the last quarter century. But while 'levelling up' agendas allude to geographic inequalities, precise geographic causes, scales of operation and interactions with individual characteristics across entire populations are not addressed. Long-run geographic effects have been considered only relative to macro characterisations of city specialisation, independent of relative location (Swinney & Thomas, 2015). Variation in life chances within as well as between both long-established and migrant groups is poorly characterised when subsumed within inherently vague categories such as 'South Asian' or 'White British' (Longley et al., 2015, pp. 103–108). Consideration of the fundamentally inter-generational origins and operation of inequalities is typically restricted to generational parent – child outcomes (Social Mobility Commission, 2020b). Transitory measures of income and welfare payments are used instead of more permanent measures such as residential neighbourhood quality that not only better manifest the social mobility outcomes in relative deprivation but can also drive increasing divergence.

Britain's population is thus characterised by inequalities of outcome that have geographic, ethnic, and familial dimensions (Brynin & Güveli, 2012; Clark & Cummins, 2014). Previous research has shown that these inequalities are formed and shaped by both long-term and short-term processes. Differences amongst long-established British family groups, for instance, are overlain by geographic factors (Longley et al., 2021), while the Covid-19 pandemic has been shown further to exacerbate socio-economic inequalities between ethnic groups (see, for example, Nafilyan et al., 2021). In this paper we analyse population-wide inequalities of outcome at the level of the individual and family group with respect to ancestral geographic origin, ethnic group, and probable country of origin. Systematic identification of the geographic origins of the longest settled groups is difficult to ascertain in the absence of systematic survey data, but we are able to use digital records from the 1851 Census of Population (Higgs & Schürer, 2019) to provide a population-wide snapshot prior to the advent of wider migration and social mobility in the 20th and 21st centuries.

We consider 'family group' alongside geography in order to identify individuals that likely share common national, regional or even local geographic origins (Winney et al., 2012). This term does not imply direct genealogical co-ancestry – albeit that this is probable amongst bearers of rare family names. The common practice of passing family names along the male line makes it possible to relate social mobility outcomes for the widest range of different family groups over truly inter-generational time periods. This process can contextualise and extend studies of changes in occupational and income characteristics between the last two generations (e.g. Dearden et al., 1997). Use of family groups can identify population-wide inequalities of outcome within as well as between long-established and recent migrant groups, without recourse to data anonymisation and denudation of geographical context (see Aspinall, 2021).

In successive stages, our population-wide analysis considers outcomes for:

- a. long-established White British populations that have differing regional origins, in the context of enduring uneven regional development¹;
- b. predominantly long-established Irish migrants² relative to migrants of more recent origin;
- c. predominantly recent migrant groups, as aggregated using a standard Office for National Statistics (ONS) classification (Office for National Statistics, 2009); and
- d. migrants from different countries subsumed within blanket 'White Other', 'Other Asian', 'Black African', 'Black Caribbean', 'Other Asian' and 'White Other' ONS classification categories.

Outcomes are defined using a harmonised measure of neighbourhood deprivation (hardship) at the individual's place of residence. We measure outcome for every family group as the mean neighbourhood deprivation experienced by each adult family group member. Through grounding of this analysis at the level of the individual, we are able to develop a more detailed taxonomy of outcomes than is possible using conventional statistical sources. This is made possible by three methodological innovations in the creation and maintenance of Personal and Sensitive Personal individual level data in a Trusted Research Environment (TRE). First, we apply name classification techniques to historical and present-day population lists to ascribe probable ethnic and nationality characteristics. Second, we geo-locate recent (2016) residential locations of all adult individuals in Great Britain to measure present day social mobility outcomes in terms of residential neighbourhood quality. This makes it possible to use a full range of individual and family group outcomes to describe inequalities between ethnic groups. Third, we use individual census records from the 1851 Census to estimate the regional origins of long-established family groups, making it possible to link the geographies of family group origins to social mobility outcomes. In our conclusions we discuss how these analyses facilitate consistent interpretation of the effects of ethnicity, family group and place in shaping social mobility outcomes in Britain today, and consider what this might mean for Britain's economic assimilation of immigrants.

2 | DATA AND METHOD

2.1 | Individual level population-wide data for historical and present-day Britain

Within Britain, as with many nations, surnames can be used to trace their regional and local origins that have endured since they were first coined between the 12th and 14th centuries (Cheshire & Longley, 2012; Kandt et al., 2020). The widespread (albeit not universal) practice of surname inheritance through the patrilineal line makes it possible to establish the persistence of long-established family groups. Here, we define 'long established' White British family groups as having at least 100 surname bearers in the earliest available digitised Census for England, Wales, and Scotland. Individual records of their historical occurrences, including residential addresses, parish of residence and birthplace are available through the Integrated Census Microdata project (I-CeM: Higgs & Schürer, 2019). Here the individual historical decennial Censuses records for England and Wales from 1851 to 1911 and for Scotland from 1851 to 1901 (1871 excluded) have been digitised and linked to the consistent parishes created by Southall (2012, 2014).

For more recent population data, we make use of the 1997–2016 Linked Consumer Registers (LCRs: Lansley et al., 2019). The LCRs comprise near complete (see Lansley et al., op cit) and contemporary lists of given names, surnames and addresses of most resident UK adults, and have been precisely georeferenced using Ordnance Survey's AddressBase Premium database. In what follows, we extract all georeferenced individuals for the year 2016 from the LCRs and attribute a present-day neighbourhood deprivation score to each (see Section 2.3).

2.2 | Characterising Britain's long-established and immigrant populations using names-based classification

Individuals' forename – surname pairings have been used successfully to infer the probable cultural, ethnic, and linguistic origins and affinities (Mateos et al., 2011), as well as the present-day country in which the name is most likely to occur (Lan & Longley, 2022). This approach has been successfully deployed in many social, medical and management science equality audits (e.g., Easton & Pryce, 2019; Lan et al., 2021; Nathan, 2015; Thomas et al., 2021). Uncertainty regarding the classification of surnames into ethnic groups is an important issue, but has been mitigated in studies which have compared estimated ethnicities with the categories with which survey respondents self-identify (Kandt & Longley, 2018). Here we use Lan and Longley's (2022) names classification because it was developed to bring particular focus to Britain's immigrant groups. This makes it possible not only to identify the most probable country of origin of present-day surname bearers, but also to determine whether their forbears were present in significant numbers in 1851. As a first step, the names classification is used to assign each adult individual in the 2016 LCR to a group in the ONS ethnicity classification used in the 2011 and 2021 Censuses (Office for National Statistics, 2009).

The design of the ONS ethnicity classification emphasises the principal origins of Britain's immigrant groups – Bangladesh, India, Ireland, Pakistan, and the 'Caribbean', but amalgamates 'Black African', 'Other Asian', and 'Other White' groups into blanket categories that likely subsume different amounts of human and social capital: some migrants

are less well educated, qualified or socially organised than others, with highly qualified economic migrants and refugees typically providing polar examples. Such differences vary systematically between individual countries of origin and are not always picked up in aggregated census type classifications. For this reason, in subsequent analysis we further disaggregate these groups, together with ‘Black Caribbeans’, according to most probable countries of origin. Established residence for the ‘White British’ group is confirmed for family names that have more than 100 bearers in 1851, with other names classified as White British fed back into the classification to identify any probable external country of origin. White Irish migrants are divided into ‘long-established’ family groups with more than 100 bearers in Great Britain in 1851 and ‘recent’ family groups classified as ‘White Irish’ but with fewer family group representatives in 1851. Figure 1 summarises the steps used to assign probable ethnic affiliations to the 2016 LCR records. The small number of individuals classified as ‘Any Other’ ethnicities are excluded from the subsequent analysis.

In what follows we variously analyse inequalities of outcome using statistics pertaining to individuals and family groups. Individuals are assigned using a system of weights applied to forenames and surnames. It is anticipated that bearers of the same family name (but different given names) may be assigned to different ethnic groups – as with the name Gill, which has different etymological origins in Britain and Pakistan, for example. We require a minimum threshold of 45 name bearers to be reached before a family name is allocated to an ethnic group.

2.3 | Attributing present-day neighbourhood circumstances to family groups

Our chosen measure of social inequality is the Index of Multiple Deprivation (IMD: Ministry of Housing, Communities, & Local Government, 2019), published at Lower layer Super Output Area (LSOA, typically 1500 residents) scale or equivalent (Data Zones in Scotland, typically 784 residents). We reconcile the 2019 scores for England and Wales with 2020 scores for Scotland by assuming the index scores are broadly comparable (see also Longley et al., 2021): although differing in detail, the national measures for constituent UK countries are broadly comparable with respect to physical and social conditions such as income, employment, health, education, crime, barriers to housing and general living environment. As such, they are akin to ‘permanent income’ measures (Clark & Moore, 1980). We range standardise the

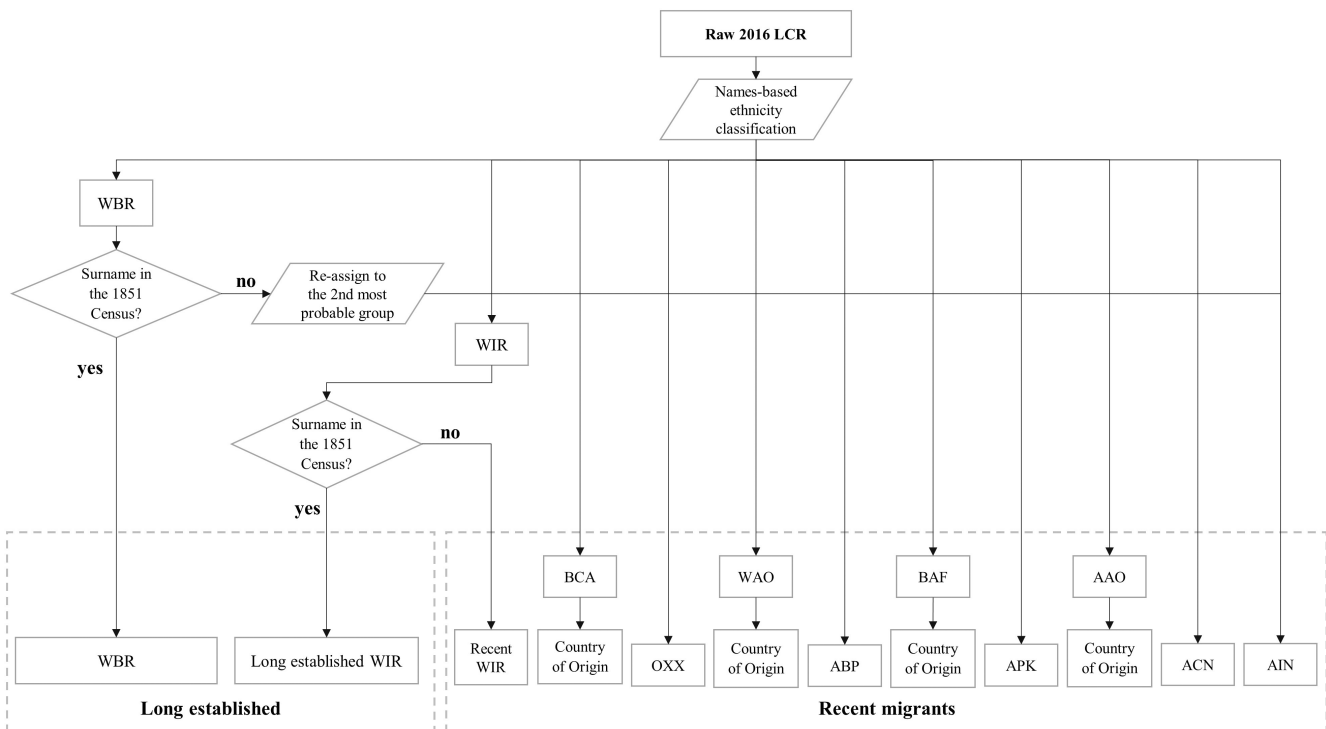


FIGURE 1 Procedures used to estimate ethnicity and country of origin of all individuals recorded in the 2016 Linked Consumer Register (AAO Asian Other; ABD Bangladeshi; ACN Chinese; AIN Indian; APK Pakistani; BAF Black African; BCA Black Caribbean; OXX Any Other; WAO White Other; WBR White British; WIR White Irish). Instances of OXX include individuals of mixed ethnicity and are excluded from the subsequent analysis.

harmonised IMD measures such that the mean individual IMD score is 50: 'high' and 'low' IMD scores may be identified relative to this benchmark figure. We assign each adult individual an IMD score based upon residence and each family group an average IMD based on all individual scores. While the quality of neighbourhood experienced by individuals and households may change over any individual's lifespan, we assume such variations average out within family groups present in Britain for at least a generation. It is also assumed that averaging takes place where spouses take the surnames of their partners, consistent with thinking about assortative mating (Clark & Cummins, 2022), albeit that some outcomes will arise from mixed ethnicity partnerships and not be correctly assigned to the OXX (Any Other ethnicities) category. Any observed variation between family group mean deprivation scores is thus deemed to relate to family group origins, either within the UK or internationally.

The contemporary social mobility outcome for each family group is defined as the average of each family name-bearer's neighbourhood deprivation score. Statistical measures of inequalities in IMD outcomes between and within groups can be calculated for (a) long-established White British or White Irish groups compared to more recent arrivals; (b) standard ONS ethnicity reporting categories; and (c) individual countries within blanket ethnic groupings. Retention of family group enables analysis of its effects within each ethnic group or national category. A logical expectation is that family group means and individual means will be identical within each category. Deviations from this indicate that family group membership compounds ethnic or national group differences, with, for example, higher family means indicating additional family group capital.

2.4 | Georeferencing the origins of long-established white British (WBR) family groups

Cheshire and Longley (2012) demonstrate that most long-established British family names (aside from relatively few common names such as Smith, Brown, and Williams) have discernible regional or even local origins, albeit that singular 'heartlands' are by no means universal. Detection of origins may also be obfuscated by the effects of rural – urban migration that had already occurred prior to the first available (Lan & Longley, 2022) digital census data. Van Dijk and Longley (2020) nevertheless establish that most rural White British bearers of regionally concentrated surnames reside in the environs of where the surnames originated. Here we use kernel density estimates to filter names that have three or fewer regional concentrations in 1851. This results in our retaining 4284 long-established names (e.g. Figure 2a) and discarding 19,473 others (e.g. Figure 2b). This is an expedient solution to a complex and intractable problem that merits further research but, imported Irish names aside, we do not find any over-all geographic bias to the names that are retained. The 1851–1891 consistent parish centroids are then assigned to residences of all bearers of retained names. Geographic centroids are calculated for each family name. Each family name centroid is then attributed with the mean IMD score of all of its adult bearers in 2016.

3 | RESULTS

3.1 | Regional outcomes for long-established white British family groups

The grand mean harmonised IMD for all adult members of the 4284 long-established White British (WBR) family groups identified in Section 2.4 is 53, with a standard deviation of 4.6 and an interquartile range of 50–57. Figure 3a shows the systematic geographic variation in family group means, with an over-all north – south divide in England (albeit with local concentration of disadvantaged family groups in the extreme southwest) and an east – west divide in Scotland. At local scale, the disadvantaged areas of these nations nevertheless include some family groups that are relatively advantaged, as well as some patterns of within-region inequality of outcome. Figure 3b shows that standard deviations of the family group means are generally highest in northern England and southern Scotland. Viewed in the context of the inter-generational southerly drift of the British population, this likely manifests the improved circumstances of the minority of disadvantaged family group members that have migrated to more prosperous areas over time. The positive skew (Figure 3c) of these disadvantaged areas (and in the southwest) confirms the preponderance of less favourable outcomes compared with the more even distribution of outcomes elsewhere. The map of kurtosis (Figure 3d) suggests more even outcomes for family groups concentrated in southern England. Geographically weighted regression (Fotheringham et al., 2003) of mean family group IMD scores against their corresponding standard deviations reveals a statistically significant negative mean parameter value of -1.19 (Table 1), confirming that family groups experiencing the highest



FIGURE 2 Kernel density estimates (KDEs) of the 1851 population-weighted geographic distributions of bearers of the family names (a) Charlton and (b) Bale. The blue contours enclose areas where bearers of the name were most concentrated, using the population-weighted density measure of Van Dijk and Longley (2020).

neighbourhood quality (best) scores tend to have the lowest variability in outcomes. The adjusted R^2 statistic and standard deviation score attest to the strength of this relationship. The wide range of geographically weighted parameter values is consistent with the patterns shown in Figure 3. (We do not add skewness or kurtosis to this regression because the variables are highly collinear).

3.2 | Comparative outcomes for long-established populations and more recent migrants

The IMD distributions of the long-established (WBR and WIR) populations and the more recent migrants are shown in Figure 4, followed by the Kruskal-Wallis H and the post hoc Dunn's tests in Table 2. Individual members of the long-established populations are on average much better off than recent migrants – the same is true for family group aggregations although there is no significant difference within the WIR group. The WBR individuals and family groups on average have higher mean and median IMD scores than the other three populations. Individual level and family group analysis indicates that the WIR are disadvantaged relative to the WBR majority, albeit by a small margin and with similar interquartile ranges. The recent WIR migrants have slightly higher mean and family group scores than their long-established compatriots. At family group level, the interquartile ranges of the Other Recent Migrants are much wider than the other three groups. Over-all, the differences between the WBR and WIR are small relative to those between the WBR and all other predominantly more recently arrived immigrants. The

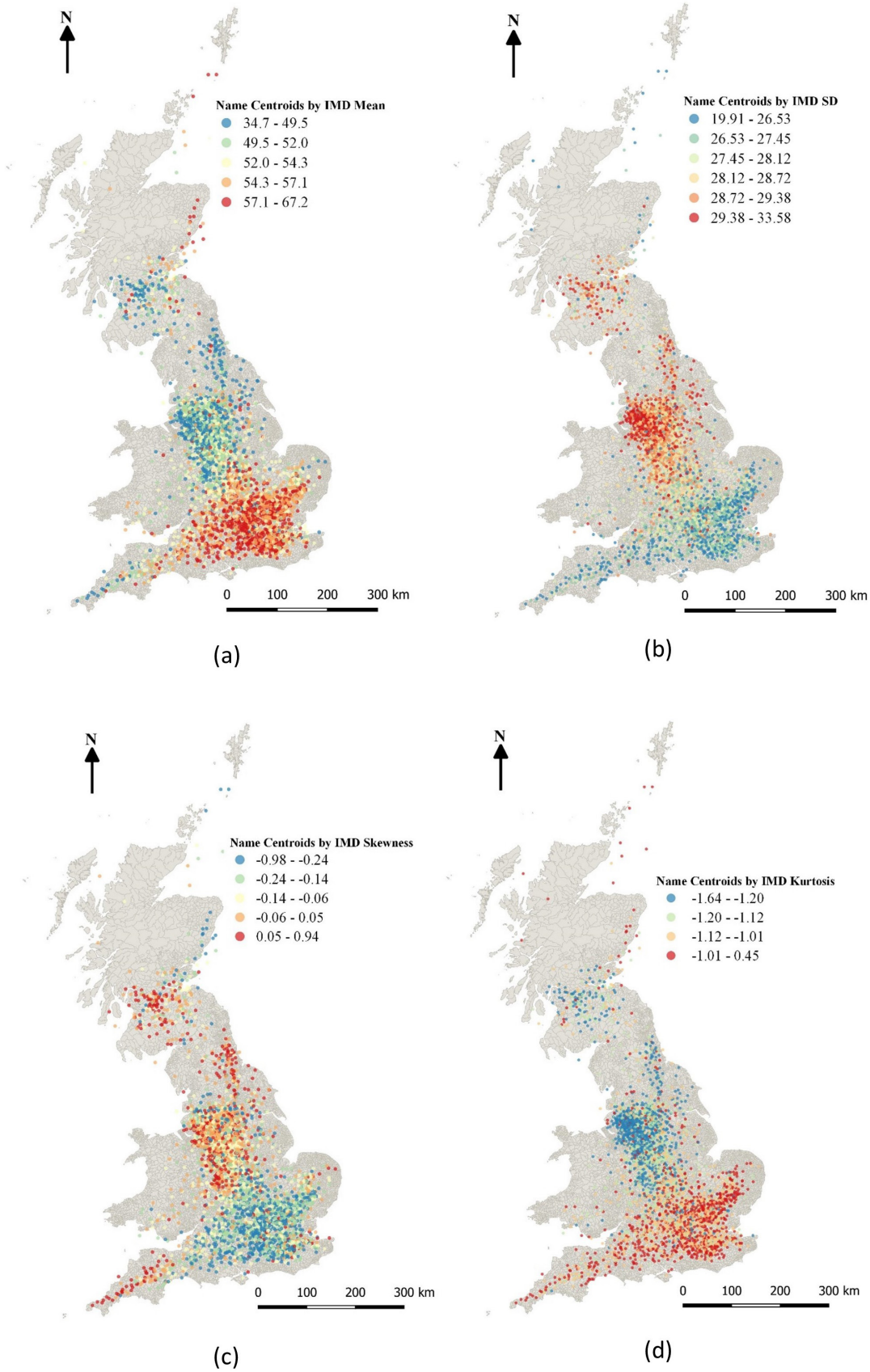


FIGURE 3 Population weighted geographic centroids of family names having three or less KDE cores in 1851, Centroids are colour coded by family group mean harmonised Index of Multiple Deprivation (IMD) (a) mean; (b) standard deviation (SD); (c) skewness; and (d) kurtosis.

TABLE 1 Summary geographically weighted regression statistics of mean family group IMD scores against corresponding standard deviations.

Parameter	Mean	STD	Min	Max
Intercept	86.64	13.03	37.06	127.39
IMD SD	-1.19	0.44	-2.63	0.37
Adjusted R ²	0.57			

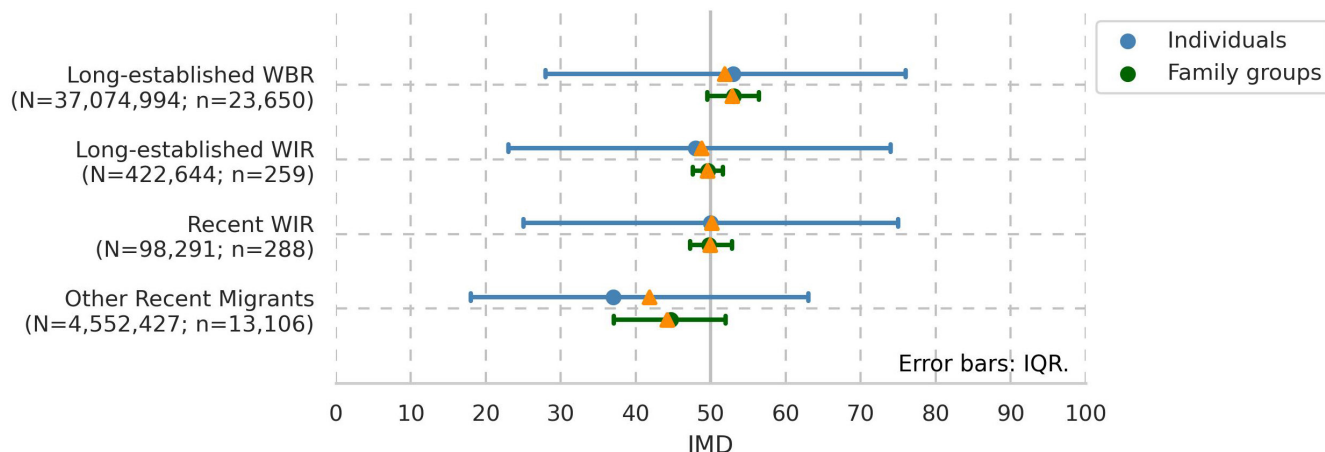


FIGURE 4 Mean (triangles) and median (circles) IMD scores and inter-quartile ranges (IQR) for individuals (upper bars) and family groups (lower bars) belonging to the White British, long-established WIR, Recent WIR and all other recent immigrant populations. *N* and *n* represent the numbers of individuals and family groups, respectively.

TABLE 2 The Kruskal–Wallis *H* and the post hoc Dunn’s tests between each pair of WBR, long-established WIR, recent WIR and other recent migrants.

	WBR	Long-established WIR	Recent WIR	Other Recent Migrants
WBR				
Long-established WIR				
Recent WIR		FALSE		
Other Recent Migrants				

Note: FALSE denotes there is no statistically significant difference, otherwise off-diagonal differences are significant by default. The highlighted lower triangle below the principal diagonal shows the statistical test results for family groups and the upper triangle shows the test results for individuals.

Abbreviations: WBR, White British; WIR, White Irish.

difference between long-established WIR individuals and more recent arrivals should be treated with caution, as the significant difference is not replicated at family group level: this may be attributable to highly variable numbers of individuals belonging to different family groups, or the inherent vagaries of assignments to the long-established and recent WIR groups.

3.3 | Differences in outcomes between conventionally defined ethnic groups

Outcomes arranged using the principal categories of the ONS ethnicity classification are shown in Figure 5. On average, individuals from the White (WBR, WIR and WAO) and Chinese (ACN) groups live in the least deprived neighbourhoods; at the other end of the spectrum, Black African (BAF), Pakistani (APK) and Bangladeshi (ABD) individuals on average

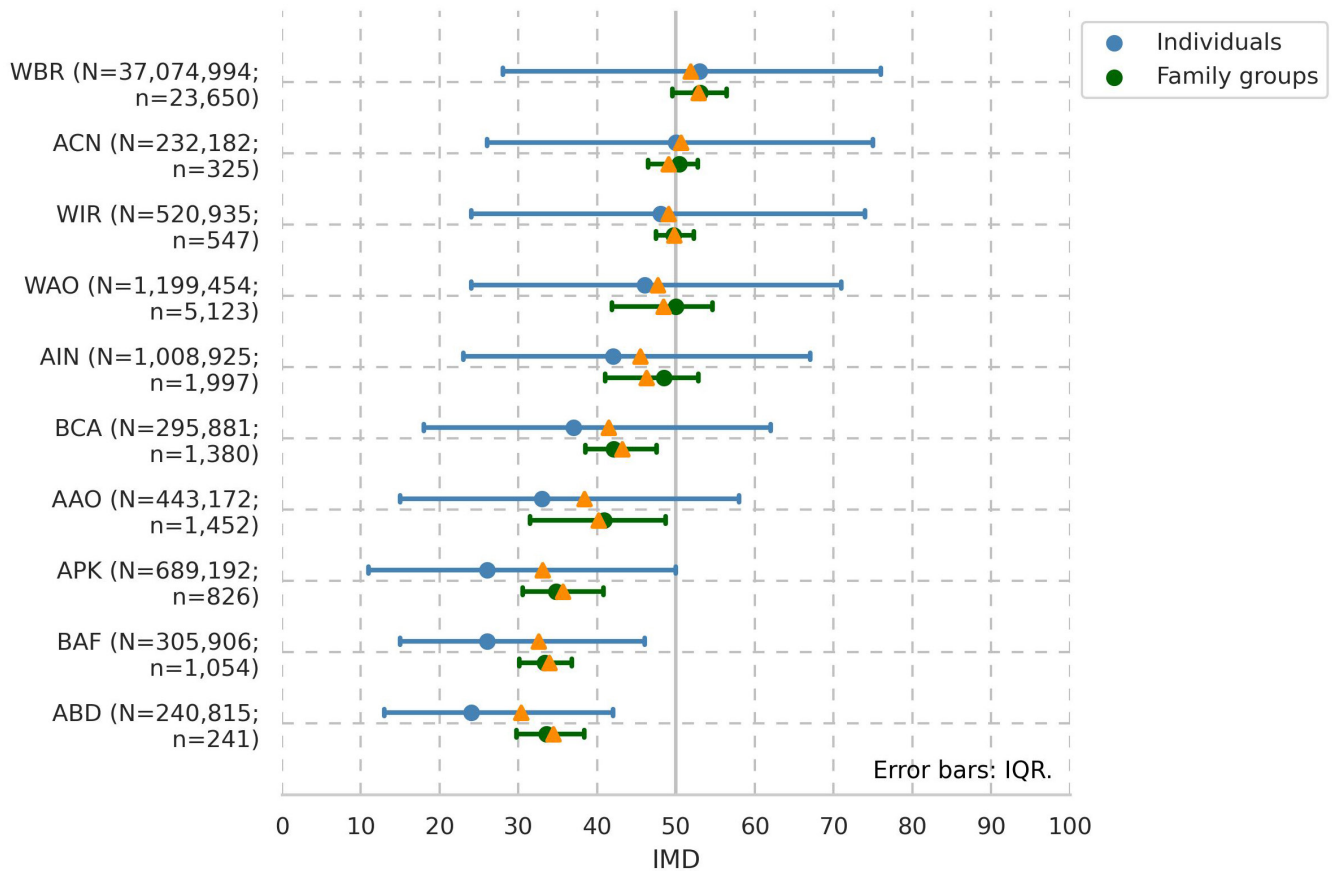


FIGURE 5 Mean (triangles) and median (circles) ethnic group IMD scores and interquartile ranges for individuals (upper bars) and family groups (lower bars) assigned to each principal ONS ethnic group. (N and n are the numbers of individuals and family groups respectively. AAO, Asian Other; ABD, Bangladeshi; ACN, Chinese; AIN, Indian; APK, Pakistani; BAF, Black African; BCA, Black Caribbean; OXX, Any Other; WAO, White Other; WBR, White British; WIR, White Irish).

experience very high levels of deprivation. Mean and median values of neighbourhood deprivation experienced by individuals within the White and Chinese groups are very similar and are approximately normally distributed (not shown). Mean and median values for individuals are more widely spaced for more recent migrant groups, and distributions of outcomes are negatively skewed towards highly deprived outcomes: these characteristics become more pronounced amongst the most deprived ethnic groups. The non-parametric Kruskal-Wallis H and pairwise post hoc Dunn's tests (Table 3) confirm that there are significant differences between the mean IMD values of every pair of ONS groups at the individual level of analysis.

Mean and median family group scores are more similar than the corresponding individual values, with the differences for the WAO and AIN groups suggesting a negative skew. Family group means are higher than individual means for the most deprived ethnic groups. An implication is that the group means at individual level are down-weighted by the neighbourhood circumstances of bearers of the more common names within the group. The ranking of mean family group scores is broadly similar to the individual level analysis, apart from the Chinese and White Irish groups, although the differences in means between ACN and each of WIR, WAO, and AIN are statistically insignificant. The family group differences in means between the Black Caribbean and Asian Other, between Bangladeshis and each of Black Caribbean, Black African, and Pakistani groups are not statistically significant.

Taken together, the individual and family results shown in Figure 5 confirm the favourable neighbourhood circumstances of the White and Chinese ethnic groups. The former groups are primarily long settled in Britain or, in the case of the White Other (WAO) group, predominantly comprise individuals from countries that are economically developed; the latter also have a long history of migration to Britain, the most recent of which are typically highly skilled.

TABLE 3 The Kruskal–Wallis *H* and the post hoc Dunn’s tests between each pair of the 10 ONS Groups (excluding OXX).

	WBR	ACN	WIR	WAO	AIN	BCA	AAO	BAF	APK	ABD
WBR	XXXXXX									
ACN		XXXXXX								
WIR		FALSE	XXXXXX							
WAO		FALSE	FALSE	XXXXXX						
AIN		FALSE			XXXXXX					
BCA						XXXXXX				
AAO						FALSE	XXXXXX			
BAF								XXXXXX		
APK									XXXXXX	
ABD								FALSE	FALSE	XXXXXX

Note: FALSE denotes there is no statistically significant difference, otherwise significant by default. The highlighted lower triangle below the principal diagonal shows family group test results and the upper triangle shows the results for individuals.

Abbreviations: AAO, Asian Other; ABD, Bangladeshi; ACN, Chinese; AIN, Indian; APK, Pakistani; BAF, Black African; BCA, Black Caribbean; ONS, Office for National Statistics; OXX, Any Other; WAO, White Other; WBR, White British; WIR, White Irish.

3.4 | Disaggregation of the AAO, BAF, BCA and WAO groups

The divergence of mean and median individual IMD scores amongst the more deprived ONS groups is also manifest within the countries making up the blanket Asian Other (AAO), Black African (BAF), Black Caribbean (BCA) and White Other (WAO) groups (Figure 6). Although having high overall average IMD scores of individuals compared to other ethnic groups in Figure 5, the WAO group has polarised extremes of Western and predominantly Eastern European countries. The AAO group results show that emigrants from more economically developed countries, including Japan and South Korea, are more likely to live in less deprived neighbourhoods. Differences amongst (the smaller number of) BCA migrant origins are less apparent. Within the BAF group, individuals that have Somalis names are amongst the lowest IMD scoring communities compared to these who emigrated from the countries with previous colonial ties with Britain or France, such as individuals from South Sudan and Benin. The individual level IMD disparities between the highest- and lowest-ranked WAO and AAO nations are greater than these within the BCA and BAF groups.³ Mean and median values are almost coincident amongst the least deprived WAO and AAO groups, but they are different, indicating negative skewness for all other nations. Table S1 in the Supplementary Information shows that there are few statistically significant differences in mean IMD scores, amongst the highest scoring countries within each ONS category.

4 | DISCUSSION

The findings of this nationwide research have important implications for the ways in which we evidence inequalities in Britain. We identify clear geographic inequalities within the long-established White British group; we demonstrate how inter-generational migration has mitigated, but not overcome, these geographic inequalities; we identify how family group effects remain much more significant for more recent migrant groups than for those much longer established; and we demonstrate differences within as well as between family groups as conventionally defined. Here we address each of these findings in turn.

The results support the use of names as tokens of regional origins and national identities that are associated with generational and inter-generational inequalities of outcome in Britain today. Coupling and georeferencing of historical and present-day name-and-address lists pertaining to entire populations make possible quantification of the effects of place upon social outcomes, while ethnicity or nationality estimation provides a framework for evaluation of the effects of national origin and family group membership. In the case of Great Britain, long-established family groups today typically live in neighbourhoods that are better than average and the variability of family group outcomes is lower than for more recent immigrant groups. Yet historical (12th–14th centuries) family group origins in the north of England or west of Scotland are typically associated with residence in more deprived neighbourhoods today – albeit that anomalous family group scores may indicate within-region patterns of intermarriage amongst family groups set apart from regional trends. The skewness and kurtosis of family group outcomes show clear spatial patterning with respect to family name centroids – indicating greater variability of within-region outcomes for bearers of names originating in disadvantaged areas. This is consistent with better outcomes for the minority of migrants from these areas (and their descendants) to more prosperous areas. Names originating in southern England and the east of Scotland have consistently better

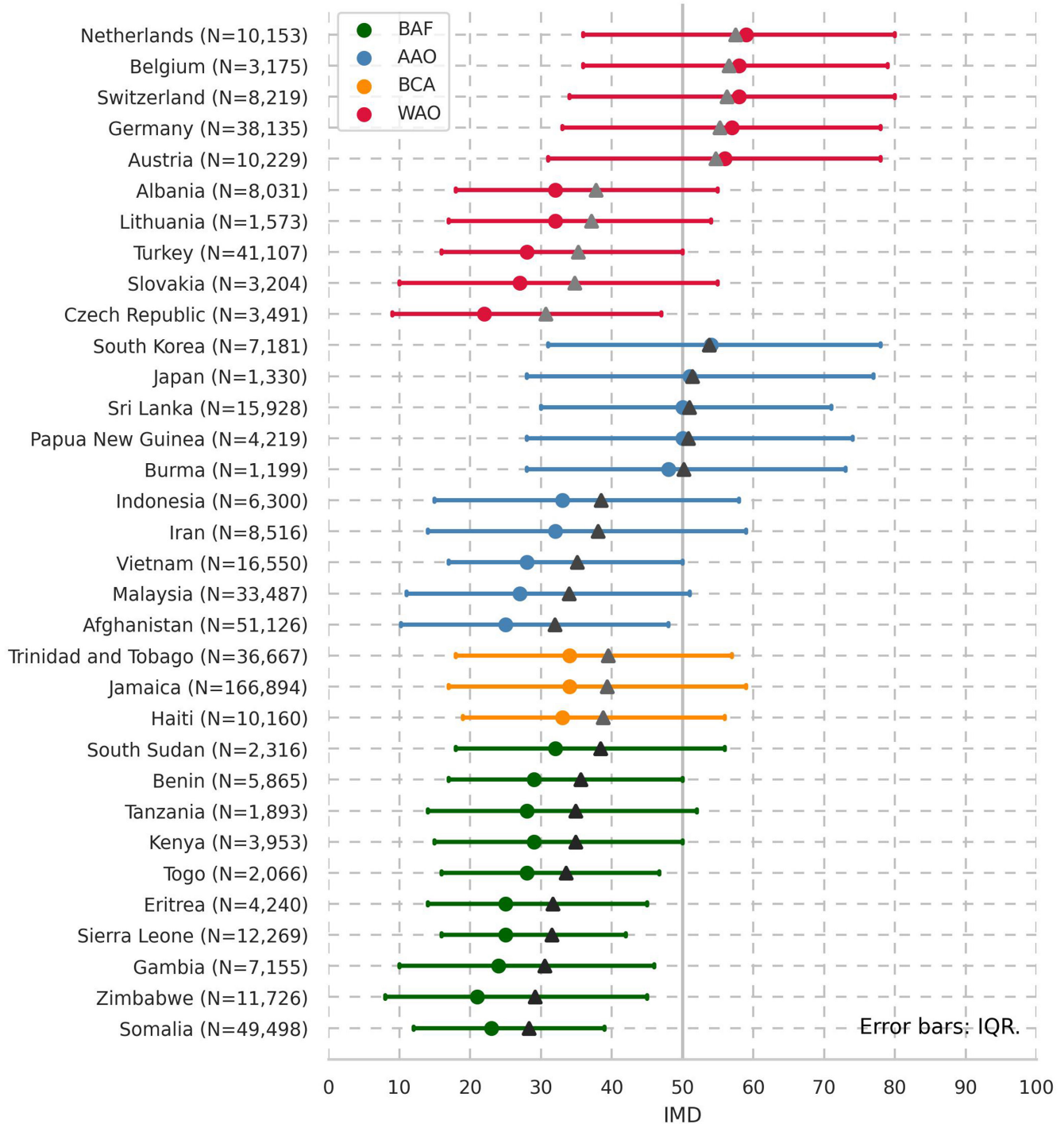


FIGURE 6 Mean (triangles) and median (circles) ethnic group IMD scores and interquartile ranges for individuals for the countries in the WAO, AAO, BCA and BAF groups with the highest and lowest mean scores. Numbers in brackets identify the numbers of individuals in each group.

outcomes although a combination of positive skew and relatively low kurtosis may indicate greater variability amongst names originating in southwest England.

The White Irish are the largest group to have sustained migration to Great Britain over centuries. Their overall mean individual and family group deprivation scores are very similar to those of the White British, with some evidence that the best outcomes accrue to members of family groups identified predominantly as later arrivals. Although our indicator of recency of White Irish migration is very imperfect, the results nevertheless suggest some lack of social mobility amongst the descendants of long-established Irish migrants.

The standard ONS ethnicity classification provides a useful framework for evaluating many of the principal present-day inequalities of outcome between different population groups. It is the case that White groups account for three of the four least deprived groups and that their neighbourhood living circumstances are far less deprived than most ethnic groups. But it is somewhat simplistic to conceive of this structure as an undifferentiated divide between White and other groups: on average, individuals of Chinese extraction enjoy similar neighbourhood circumstances to White groups, while differences between all other ethnic groups are statistically significant. This interpolation is also true at family group level, with the exceptions of differences between AAO and BCA, between ABD and APK, and between ABD and BAF.

Family group and individual level means and medians are very similar to one another for the long-established White British, as well as for what may be a mix of long-established low-skilled and more recent high-skilled Chinese and White Irish immigrants. The interquartile ranges and near coincidence of mean and median values suggest that family group membership is less important for these groups relative to other, typically more recent, immigrant groups. With the exception of the Indian group, higher family group medians suggest that relatively common family names depress individual mean and median scores and that, to varying degrees, membership of some rarer family groups is associated with less deprived outcomes. This effect becomes progressively more pronounced amongst the most disadvantaged ethnic groups. The successive ordering of other non-White groups is: Indian, Black Caribbean, Other Asian, Pakistani, Black African and Bangladeshi, and differences in individuals' mean deprivation are all statistically significant. Individual outcomes are positively skewed within all groups: this outcome is highest amongst the most deprived ethnicities.

Analysis using the ONS classification identifies differences in outcome within the Indian sub-continent, although this is not apparent for Africa, other parts of Asia, the Caribbean and the diverse national origins of White migrants. Analysis of constituent countries indicates significant within-group differences throughout all but the Caribbean group. The range of outcomes is greatest within the White Other and Asian Other groups, in which the least deprived nationalities attain deprivation scores that are mostly significantly better than those recorded for the White British, Chinese and White Irish groups. This finding may manifest both greater migrant selectivity, ambient human and social capital in the origin countries and inter-generational length of establishment in Great Britain. In the case of the White Other and Asian Other groups there is positive skewness only amongst individuals originating in the worst performing nations: the distributions of most African and Caribbean nations are positively skewed towards more deprived outcomes.

Taken together, these results provide evidence that present-day outcomes for White Irish and Chinese immigrants are similar to those for the long-established White British, consistent with a mix of long-established but non-selective migration and more recent selective migration. Family group membership is unimportant for these groups. More recent migrant groups predominantly endure much inferior neighbourhood conditions and family group membership is associated with differing within-group outcomes. The least deprived White Other and Asian Other nationalities resemble the White British, Chinese and White Irish in terms of mean deprivation scores and absence of family group effects, while family group effects are present amongst the most deprived nationalities in these groups. There are overlaps between these groups – for example, the most deprived White Other nations are more deprived than the least deprived classed as Asian Other, with similar outcomes for the most deprived Asian Other and least deprived Black African nations. Blanket terms describing ethnic minorities mask such within-group variations.

5 | CONCLUSIONS

Great Britain received an estimated 8,570,000 migrants between 1800 and 2010, of which some 73% (6,231,000) arrived between 1945 and 2014 (Panayi, 2014). More recently the UK received c. 513,000 long-term international migrants in 2021 (Office for National Statistics, 2022) under schemes including asylum claims, responses to labour market shortages and selection of highly qualified graduates. Diversity of personal circumstances brings the implication that the 16.8% of residents of England and Wales identified by the 2021 Census as born outside the U.K. will have likely brought very different stocks of social and human capital.

The quantitative analysis developed in this paper makes methodological contributions to measuring, and hence understanding, the nature and dynamics of ethnic diversity – as demonstrated here for Great Britain but also internationally. The literature on residential segregation frequently acknowledges that reliance upon aggregate data limits the scale range over which the phenomenon may be measured, and cautions about the attendant potential risks of ecological fallacy (see Barros & Feitosa, 2018). The approach developed here enables modelling ethnic classification at the most appropriate level – that of the human individual, georeferenced to very high precision. As a result, scale-free representations of ethnic distributions may be developed for linkage, in secure environments, to other micro

representations of living or other personal circumstances. Individual occurrences may also be linked to aggregate data pertaining to neighbourhood conditions (Jivraj & Simpson, 2015; Johnston et al., 2016). In such instances ecological fallacy may arise in unusual specific instances – it may be the case, for example, that members of an ethnic minority predominate in the worst, or best, housing within one particular Lower layer Super Output Area (LSOA) – but the magnitudes of the differences identified here render this extremely unlikely across an entire national system of LSOAs and Data Zones.

A further contribution of this paper is to stimulate reconsideration of ethnicity, family group and ancestry as determinants of economic inequalities of outcome in Britain and other advanced societies. A central motivation for spatial analysis of ethnicity has been as an indirect indicator of economic assimilation into host societies (e.g. Dubuc, 2012; Finney, 2011). Our analysis suggests that the outcomes of such assimilation require inter-generational perspectives for two reasons: first, because this enables outcomes to be analysed for ethnic groups with different typical multigenerational residence durations (as here for the White Irish, for example); and second, because the accident or choice of residential location can affect the outcome (as with the case of the long-established White British group). These effects can be monitored over any convenient timescale (see also Catney et al., 2020) for which names data are available, and are not limited to census data availability (cf. Catney, 2016, 2018). Assessment of inter-generational outcomes is facilitated by coupling of surnames data and can accommodate inter-generational transfers of wealth or stocks of social and human capital (e.g., Adermon et al., 2021; Clark & Cummins, 2014).

The use of family names also reveals that conventional aggregations, such as those used in official statistics, may subsume considerable variation in outcomes (as observed here amongst the White Other and Other Asian categories: see also Aspinall, 2000). The evidence of our analysis suggests that outcomes for long-established White British family groups remain geographically differentiated but, overall, show much less variation than those attributed to more recently arrived family groups. Although most recently arrived groups endure worse (sometimes very much worse) neighbourhood circumstances than their long-established peers, positive skewness of family group outcomes amongst recent migrant groups suggests positive benefits of more advantaged family groups. Indeed, the interquartile ranges of the White British family groups have significant overlaps with all but one White Other, one Other Asian and six Black African national groups. However, interquartile ranges of family group means for Bangladeshis and Pakistanis are very much lower than for the White British as, when aggregated, is the range for Black Caribbeans, Black Africans and the Asian Other categories. This suggests that blanket descriptions such as BAME ('Black, Asian and minority ethnic') may be unhelpful in equality audits, and that individuals may identify with their family group outcome rather than that of their ethnic group: the mean standardised IMD score for the McFee (White British) family group is 44.83, for example, while that of the Banerjees (Indian) is 53.69.

This work has several implications for understanding economic assimilation of migrants and the persistence of inequalities amongst and between populations. The immobility of the long-settled British population means that the majority of region- and locality-specific surnames remain concentrated in the areas in which their names were first coined some 700 or more years ago. The opportunities available to this majority remain heavily constrained by these accidents of history. Inter-generational effects have not been entirely overcome by migration across the generations. The variability of outcomes between family groups is nevertheless small relative to that observed amongst more recent migrant groups. Most of these, including the long-established Irish, are disadvantaged relative to the White British but differences between family group and individual outcomes suggest family group membership retains important human and social capital for some recent migrants.

These results invite many policy-relevant extensions and developments. First, the source data are one annual snapshot from a set of 'Linked Consumer Registers' (Lansley et al., 2019) and the analysis set out here could be replicated to examine trends for each group over the entire period covered by this data source. This would make it possible better to understand whether migrant assimilation is on a convergent path, and the extent to which this pathway is contingent upon British regional geography. Second, the linkage of these registers makes it possible to examine the dynamics of residential movements of different groups, in terms of resultant changes in deprivation, viewed in local and regional contexts (Longley et al., 1991; Van Dijk et al., 2021). Third, the nature of the groups might be further disintegrated through evaluation of linkage of resident names to the most likely periods in British history in which group migrations have taken place. The analysis of Irish names set out here might be extended to investigate the effects of caste or regional origins amongst immigrants.

Finally, the work has the potential to reinvigorate many aspects of quantitative geodemographic analysis of neighbourhood structure. Neighbourhoods might be characterised by their inter-generational as well as ethnic diversity; new sources and forms of consumer data might be used to characterise the activity patterns of different groups; and alignment

of consumer registers with census data might chart the consequences of migration for both long-established and more recently arrived groups in the context of 'levelling up' agendas. These approaches will each be the subject of our future investigations.

ACKNOWLEDGEMENTS

Historical census data from the I-CeM (Integrated Census Microdata) project were licensed and provided by the UK Data Service (UKDS). Linked Consumer Registers data were provided by the ESRC Consumer Data Research Centre (CDRC).

DATA AVAILABILITY STATEMENT

The 1997–2016 Linked Consumer Registers data (<https://data.cdrc.ac.uk/dataset/linked-consumer-registers>) used in this research are available upon successful application to the Consumer Data Research Centre (CDRC: <http://www.cdrc.ac.uk>). The I-CeM historical census data 1851–1911 (<https://doi.org/10.5255/UKDA-SN-7856-2>) are available from the UK Data Service (UKDS: ukdataservice.ac.uk). Restrictions apply to the availability of these data, which were used under licence for this study. The harmonised IMD data are derived from public domain resources and are available from the CDRC data repository (<https://data.cdrc.ac.uk/dataset/index-multiple-deprivation-imd>). The analysis conducted here was undertaken under special licensing arrangements for access to individual level data. As such, the data are only available upon successful application to these Economic and Social Research Council-funded data centres.

ORCID

Paul Longley  <https://orcid.org/0000-0002-4727-6384>

Tian Lan  <https://orcid.org/0000-0002-3961-4742>

ENDNOTES

- ¹ The idea of the 'levelling up' agenda refers to a series of White Papers and government strategies with the intention to spread (economic) opportunities more equally across the United Kingdom (UK Government, 2022).
- ² Until around the 1950s, the Irish were the largest ethnic minority in Britain following large successive waves of migration since the 18th century (see, for instance, MacRaild, 1999).
- ³ We do not include family group analysis because they are small in number for some countries.

REFERENCES

- Adermon, A., Lindahl, M. & Palme, M. (2021) Dynastic human capital, inequality, and intergenerational mobility. *American Economic Review*, 111(5), 1523–1548. Available from: <https://doi.org/10.1257/aer.20190553>
- Almeida, A. (2021) *Pushed to the margins: A quantitative analysis of gentrification in London in the 2010s*. London, UK: Runnymede Trust.
- Aspinall, P. (2000) The challenges of measuring the ethno-cultural diversity of Britain in the new millennium. *Policy and Politics*, 28(1), 109–118. Available from: <https://doi.org/10.1332/0305573002500866>
- Aspinall, P.J. (2021) BAME (black, Asian and minority ethnic): The 'new normal' in collective terminology. *Journal of Epidemiology and Community Health*, 75(2), 107. Available from: <https://doi.org/10.1136/jech-2020-215504>
- Barros, J. & Feitosa, F.F. (2018) Uneven geographies: Exploring the sensitivity of spatial indices of residential segregation. *Environment and Planning B: Urban Analytics and City Science*, 45(6), 1073–1089. Available from: <https://doi.org/10.1177/2399808318760572>
- Brynin, M. & Güveli, A. (2012) Understanding the ethnic pay gap in Britain. *Work, Employment and Society*, 26(4), 574–587. Available from: <https://doi.org/10.1177/0950017012445095>
- Catney, G. (2016) Exploring a decade of small area ethnic (de-)segregation in England and Wales. *Urban Studies*, 53(8), 1691–1709. Available from: <https://doi.org/10.1177/0042098015576855>
- Catney, G. (2018) The complex local geographies of ethnic residential segregation: Using spatial and local segregation measures to explore scale-dependency and spatial relationships. *Transactions of the Institute of British Geographers*, 43(1), 137–152. Available from: <https://doi.org/10.1111/tran.12209>
- Catney, G., Wright, R. & Ellis, M. (2020) The evolution and stability of multi-ethnic residential neighbourhoods in England. *Transactions of the Institute of British Geographers*, 46(2), 330–346. Available from: <https://doi.org/10.1111/tran.12416>
- Cheshire, J.A. & Longley, P.A. (2012) Identifying spatial concentrations of surnames. *International Journal of Geographical Information Science*, 26(2), 309–325. Available from: <https://doi.org/10.1080/13658816.2011.591291>
- Clark, G. & Cummins, N. (2014) Surnames and social mobility in England, 1170–2012. *Human Nature*, 25, 517–537. Available from: <https://doi.org/10.1007/s12110-014-9219-y>

- Clark, G. & Cummins, N. (2022) *Assortative mating and the industrial revolution: England, 1754–2021*. CEPR Discussion Paper No. DP17074. Available from SSRN: <https://ssrn.com/abstract=4069865>
- Clark, W. & Moore, E. (1980) *Residential mobility and public policy*. Beverly Hills, CA: Sage Publications.
- Dearden, L., Machin, S. & Reed, H. (1997) Intergenerational mobility in Britain. *The Economic Journal*, 107(440), 47–66. Available from: <https://doi.org/10.1111/1468-0297.00141>
- Dubuc, S. (2012) Immigration to the UK from high-fertility countries: Intergenerational adaptation and fertility convergence. *Population and Development Review*, 38(2), 353–368. Available from: <https://doi.org/10.1111/j.1728-4457.2012.00496.x>
- Easton, S. & Pryce, G. (2019) Not so welcome here? Modelling the impact of ethnic in-movers on the length of stay of home-owners in micro-neighbourhoods. *Urban Studies*, 56(14), 2847–2862. Available from: <https://doi.org/10.1177/0042098018822615>
- Finney, N. (2011) Understanding ethnic differences in the migration of young adults within Britain from a lifecourse perspective. *Transactions of the Institute of British Geographers*, 36(3), 455–470. Available from: <https://doi.org/10.1111/j.1475-5661.2011.00426.x>
- Fotheringham, A.S., Brunson, C. & Charlton, M. (2003) *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester, UK: John Wiley & Sons.
- Harvey, D. (2010) *The enigma of capital and the crises of capitalism*. Oxford, UK: Oxford University Press.
- Higgs, E. & Schürer, K. (2019) *Data from integrated census microdata (I-CeM), 1851–1911*. UK Data Service. SN 7856. Available from: <https://doi.org/10.5255/UKDA-SN-7856-2> [Accessed 10th April 2014].
- Jivraj, S. & Simpson, L. (Eds.). (2015) *Ethnic identity and inequalities in Britain: The dynamics of diversity*. Bristol, UK: Policy Press.
- Johnston, R., Jones, K., Manley, D. & Owen, D. (2016) Macro-scale stability with micro-scale diversity: Modelling changing ethnic minority residential segregation—London 2001–2011. *Transactions of the Institute of British Geographers*, 41(4), 389–402. Available from: <https://doi.org/10.1111/tran.12142>
- Kandt, J. & Longley, P.A. (2018) Ethnicity estimation using family naming practices. *PLoS One*, 13(8), e0201774. Available from: <https://doi.org/10.1371/journal.pone.0201774>
- Kandt, J., Van Dijk, J.T. & Longley, P.A. (2020) Family names origins and inter-generational demographic change in Great Britain. *Annals of the American Association of Geographers*, 110(6), 1726–1742. Available from: <https://doi.org/10.1080/24694452.2020.1717328>
- Lan, T., Kandt, J. & Longley, P.A. (2021) Measuring the changing pattern of ethnic segregation in England and Wales with consumer registers. *Environment and Planning B: Urban Analytics and City Science*, 48(6), 1591–1608. Available from: <https://doi.org/10.1177/2399808320919774>
- Lan, T. & Longley, P.A. (2022) An individual level method for improved estimation of ethnic characteristics. *International Regional Science Review*, 46, 328–353. Available from: <https://doi.org/10.1177/01600176221116568>
- Lansley, G., Li, W. & Longley, P.A. (2019) Creating a linked consumer register for granular demographic analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(4), 1587–1605. Available from: <https://doi.org/10.1111/rssa.12476>
- Longley, P.A., Clarke, M. & Williams, H.C.W.L. (1991) Housing careers, asset accumulation and subsidies to owner occupiers—A microsimulation. *Housing Studies*, 6(1), 57–69. Available from: <https://doi.org/10.1080/02673039108720697>
- Longley, P.A., Goodchild, M.F., Maguire, D.J. & Rhind, D.W. (2015) *Geographic information science and systems*, 4th edition. Hoboken, NJ: Wiley.
- Longley, P.A., Van Dijk, J.T. & Lan, T. (2021) The geography of intergenerational social mobility in Britain. *Nature Communications*, 12, 6050. Available from: <https://doi.org/10.1038/s41467-021-26185-z>
- MacRaild, D.M. (1999) *Irish migrants in modern Britain, 1750–1922*. London, UK: Red Globe Press. Available from: <https://doi.org/10.1007/978-1-349-27344-7>
- Mateos, P., Longley, P.A. & O'Sullivan, D. (2011) Ethnicity and population structure in personal naming networks. *PLoS One*, 6(9), e22943. Available from: <https://doi.org/10.1371/journal.pone.0022943>
- Ministry of Housing, Communities & Local Government. (2019) *English indices of deprivation 2019*. Available from: <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019> [Accessed 12th January 2022].
- Nafilyan, V., Islam, N., Mathur, R., Ayoubkhani, D., Banerjee, A., Glickman, M. et al. (2021) Ethnic differences in COVID-19 mortality during the first two waves of the coronavirus pandemic: A nationwide cohort study of 29 million adults in England. *European Journal of Epidemiology*, 36, 605–617. Available from: <https://doi.org/10.1007/s10654-021-00765-1>
- Nathan, M. (2015) Same difference? Minority ethnic inventors, diversity and innovation in the UK. *Journal of Economic Geography*, 15(1), 129–168. Available from: <https://doi.org/10.1093/jeg/lbu006>
- Office for National Statistics. (2009) *Final recommended questions for the 2011 Census in England and Wales: Ethnic group*. Available from: https://www.ons.gov.uk/file?uri=/census/2011census/howourcensusworks/howweplannedthe2011census/questionnairedevelopment/finalisingthe2011questionnaire/final-recommended-questions-2011-ethnic-group_tcm77-183998.pdf [Accessed 12th January 2022].
- Office for National Statistics. (2022) *Long-term international immigration, emigration and net migration flows, provisional*. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/datasets/longterminternationalimmigrationemigrationandnetmigrationflowsprovisional> [Accessed 21th June 2022].
- Panayi, P. (2014) *An immigration history of Britain: Multicultural racism since 1800*. London, UK: Routledge.
- Piketty, T. (2014) *Capital in the twenty-first century*. Cambridge, MA: Harvard University Press.
- Social Mobility Commission. (2019) *State of the Nation 2018–19: Social mobility in Great Britain*. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/798404/SMC_State_of_the_Nation_Report_2018-19.pdf [Accessed 14th July 2022].

- Social Mobility Commission. (2020a) *Public attitudes to social mobility in the UK, 2019–20*. London, UK: Social Mobility Commission.
- Social Mobility Commission. (2020b) *Monitoring social mobility 2013–2020: Is the government delivering on our recommendations?* London, UK: Social Mobility Commission.
- Southall, H. (2012) Rebuilding the Great Britain historical GIS, part 2: A geo-spatial ontology of administrative units. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 45(3), 119–134. Available from: <https://doi.org/10.1080/01615440.2012.664101>
- Southall, H. (2014) Rebuilding the Great Britain historical GIS, part 3: Integrating qualitative content for a sense of place. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 47(1), 31–44. Available from: <https://doi.org/10.1080/01615440.2013.847774>
- Swinney, P. & Thomas, E. (2015) *A century of cities: Urban economic change since 1911*. London, UK: Centre for Cities.
- Thomas, D., Orife, O., Plimmer, A., Williams, C., Karani, G., Evans, M.R. et al. (2021) Ethnic variation in outcome of people hospitalised during the first COVID-19 epidemic wave in Wales (UK): An analysis of national surveillance data using Onomap, a name-based ethnicity classification tool. *BMJ Open*, 11(8), e048335. Available from: <https://doi.org/10.1136/bmjopen-2020-048335>
- UK Government. (2022) *Levelling up the United Kingdom*. Available from: <https://www.gov.uk/government/publications/levelling-up-the-united-kingdom> [Accessed 8th November 2022].
- Van Dijk, J.T., Lansley, G. & Longley, P.A. (2021) Using linked consumer registers to estimate residential moves in the United Kingdom. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 184(4), 1452–1474. Available from: <https://doi.org/10.1111/rssa.12713>
- Van Dijk, J.T. & Longley, P.A. (2020) Interactive display of surnames distributions in historic and contemporary Great Britain. *Journal of Maps*, 16(1), 68–76. Available from: <https://doi.org/10.1080/17445647.2020.1746418>
- Winney, B., Boumertit, A., Day, T., Davison, D., Echeta, C., Evseeva, I. et al. (2012) People of the British Isles: Preliminary analysis of genotypes and surnames in a UK control population. *European Journal of Human Genetics*, 20, 203–210. Available from: <https://doi.org/10.1038/ejhg.2011.127>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Table S1. The Kruskal–Wallis H and the post hoc Dunn's tests between each pair of the selected countries from the WAO, AAO, BCA and BAF groups. FALSE denotes there is no statistically significant difference, otherwise significant by default. The upper triangle shows the test results for individuals.

How to cite this article: Longley, P., Lan, T. & van Dijk, J. (2023) Geography, ethnicity, genealogy and inter-generational social inequality in Great Britain. *Transactions of the Institute of British Geographers*, 00, 1–16. Available from: <https://doi.org/10.1111/tran.12622>