An economic topology of the Brexit vote

Simon Rudkin, Lucy Barros, Paweł Dłotko & Wanling Qiu

View supplementary material ⟳

Published online: 19 May 2023.

Submit your article to this journal ⟳

Article views: 2268

View related articles ⟳

View Crossmark data ⟳

Citing articles: 1 View citing articles ⟳

Routledge
Taylor & Francis Group

RSA Regional Studies Association

An economic topology of the Brexit vote

Simon Rudkin[a]* , Lucy Barros[a] , Paweł Dłotko[b] and Wanling Qiu[c]**

ABSTRACT
A desire to understand the UK voting to leave the European Union continually attracts attention. We generate a multidimensional map of the economic geography of Brexit voting at the regional level, visualising hitherto unidentified insights into the regional manifestation of leave voting. While we find broad patterns consistent with national heterogeneities and the geographies of discontent, we also demonstrate support for Brexit locates in a far more homogenous set of regions than support for remaining in the European Union. Our conclusions apply at the constituency and local authority levels and are robust to inclusion of additional cultural and economic regional characteristics.

KEYWORDS
topological data analysis; geographies of discontent; Brexit; local demographics; interaction effects; voting behaviour

1. INTRODUCTION

Evaluation of the decision of the UK to leave the European Union (EU), 'Brexit',[1] has been rooted in geographies of discontent (Dijkstra et al., 2020; McCann, 2020) and the notion that regions were being 'left behind' (Heath & Goodwin, 2017). Behind these concepts sit mixtures of demographic characteristics such as education, home ownership, social status and age which have been variously associated with Leave voting. This paper takes the multidimensional dataset highlighted in existing work on Brexit to offer new evidence from the joint distribution of regional characteristics. Specifically, we take an extensive set of parliamentary constituency socio-demographic metrics and map estimated Leave vote percentages (Hanretty, 2017) onto the joint distribution. The homogeneity of Leave voting constituencies in comparison with the Remain voting areas is demonstrated for the first time. While our results point to national identity variations, we confirm the importance of economic geography for understanding Leave voting patterns. Our contribution is thus a robust representation of the economic geography of Brexit which not only forwards understanding but also allows the direction of the post-Brexit policy response.

What follows is based upon observed proportions of key demographic characteristics within parliamentary constituencies. Motivation for studying regional aggregation lies in the fact that voting behaviour cannot be divorced from local context; the characteristics of individuals interact with the broader features of their neighbours and neighbourhood (e.g., Johnston et al., 2004; Abreu & Öner, 2020). Constituency-level aggregation allows linkage directly to parliamentary election results that comment on voter party-political allegiance. Direct parallels are drawn between the 2016 Referendum and 2019 General Election, consistent with the idea that the former shaped narratives and the subsequent 'Levelling Up' policy 'mandate' emerging from the latter.[2] Although the Brexit referendum was counted at the local authority level, constituencies are also more representative of the within-local authority variation in demographics (Hanretty, 2017). Empirical robustness checks demonstrate that the comparative homogeneity of Leave holds at the local authority level, and when alternative means are deployed to aggregate votes to constituency level. Consistently, Remain voting regions are dispersed, whilst Leave voting regions are co-located in the regional characteristic space.

CONTACT Simon Rudkin ✉ simon.rudkin@manchester.ac.uk
[a] Economics Department, Swansea University, Swansea, UK
*Present affiliation: Social Statistics Department, School of Social Sciences, University of Manchester, Manchester, UK
[b] Dioscuri Centre in Topological Data Analysis, Polish Academy of Sciences, Warsaw, Poland
[c] School of Management, University of Liverpool, Liverpool, UK
**Present affiliation: University of Exeter Business School, University of Exeter, Exeter, UK
Supplemental data for this article can be accessed online at https://doi.org/10.1080/00343404.2023.2204123

The aim of this paper is to obtain an intuitive and model-free representation of referendum voting behaviour. Our work overcomes critiques of attempts to apply statistical models to sets of characteristics which are necessarily correlated, and numerous, relative to the number of observations available for analysis. Further, abstraction from a causal model also removes any imposition of a one-size-fits-all set of coefficients upon relationships which the literature otherwise identifies as being variant across regions. For example, Becker et al.'s (2017) conclusion that there is no statistical difference between coefficients in a Scotland-only model relative to a full UK model misses the distinct character of Scottish voting behaviour documented elsewhere (Clarke & Whittaker, 2016). In mapping the joint characteristic space, we show directly how different Scottish outcomes are relative to similar English or Welsh regions.[3] Construction of our data map makes use of the Topological Data Analysis BallMapper (TDABM) algorithm of Dłotko (2019a). Primary advantages of the TDABM approach to the visualisation of multidimensional datasets are the ease of interpretability and the stability of the resulting maps to perturbations of the single parameter of the algorithm.[4] While this paper represents the first application of TDABM in regional science, TDABM has been usefully applied in the context of credit risk (Qiu et al., 2020) and stock returns (Dłotko et al., 2021).

The remainder of the paper is organised as follows. Section 2 provides a brief overview of existing empirical work relating socio-economic factors to the referendum vote at various levels of spatial aggregation. Section 3 presents the main constituency level data used in our analyses, including two-sample t-tests to understand how values of these variables differ in means between Leave and Remain constituencies. The TDABM algorithm is introduced in section 4, while section 5 constructs and analyses a TDABM plot based on chosen socio-economic axes. Section 6 considers the redrawing of the UK political map in elections after Brexit, and a range of robustness analyses for the choice of socio-demographic characteristics, the imputation method of the constituency Leave vote and the level of spatial aggregation. Section 7 discusses conclusions from the TDABM analysis in light of existing empirical work on the UK 2016 European Union (EU) Referendum result. Section 8 concludes.

2. SOCIO-ECONOMIC PATTERNS UNDERLYING THE LEAVE VOTE

Differentials between Leave and Remain support are noted across nations and at the NUTS-1 level. Scotland and Northern Ireland had stronger support for Remain than did England and Wales (Clarke & Whittaker, 2016). Likewise London was more pro-Remain than many of its neighbouring rural areas (Goodwin & Heath, 2016b). However, a body of work emphasises the role of economic geography in explaining observed patterns in voting behaviour at the regional level (Los et al., 2017). Support for the consideration of economic variation across space is provided by Billing et al. (2019), McCann

(2020) and Henderson et al. (2021), amongst many others. For every broad association of economic characteristics with Leave voting, there are notable exceptions. For example, the general correlation of low pay with Leave voting is challenged by the strong Remain support in low pay areas of Scotland (Clarke & Whittaker, 2016). Evidence from Henderson et al. (2021) combines the demographics of age, sex, education and income with national identity to underline the heterogeneity of voting behaviour at the national level. Digging into the maps of the socio-economic characteristics of parliamentary constituencies, our work also shows directly the differences between constituencies in different nations with similar characteristics. At the aggregate level we highlight a homogeneity among the constituencies voting Leave, but confirm the divergence of Scottish regions from English constituencies with identical characteristics.

The idea that place matters at a lower level of spatial aggregation is fundamental to our approach. While Goodwin and Heath (2016a) conclude broadly that Brexit voters were the 'left behind', poverty and educational inequality being key factors, they report a strong interaction between an individual's education level and the general skill level of their community: graduates in low-average-skill communities were more likely to vote Leave than graduates from high-average-skill communities. Lee et al. (2018) find the role of residential immobility – the 'Somewheres' of Goodhart (2016) – in Leave-voting to depend on area-level features (relative economic decline, recent change in immigration), and Gordon (2018) finds not only a non-monotonic effect of qualifications and of social status, but also that the area-level occupational mix has a contextual effect on individual voting propensity 'and quite probably the attitudes underlying this behaviour' (p. 103). This presents a clear rationale for the joint consideration of multiple regional characteristics. Abreu and Öner (2020) also indicate that cultural considerations operate at a subnational level, exploring constituency level differentials in attitudes on questions of redistribution, immigration and sexuality. This paper takes an economic geography perspective, demonstrating in robustness checks that our results extend to the inclusion of Abreu and Öner's social attitudes data.

Where lower level aggregations are more informative on the local context of Brexit voting, we group constituencies on similarities in their socio-economic characteristics for collective interpretation. The methodology employed in this paper achieves an intuitive combination of those spatial units with values for the socio-economic variables which are within a fixed range of each other. Once a set of constituencies with similar demographics is highlighted, the role of regional identity may be understood through comparison of peers within that set. While comparisons between otherwise similar regions is the rationale for using matching methods in Abreu and Öner (2020), the TDABM approach measures the joint distance between data points as set onto multiple axes. Where that distance is below a threshold, TDABM considers constituencies similar in precisely the way we would if trying to identify groups by eye. The key departure is that the

visualisation resulting from TDABM captures the full joint distribution of all the characteristics in an abstract two-dimensional manner. The graphs we present can be intuitively understood as aggregations of smaller level geographies with similar demographic characteristics.

Use of a data visualisation approach to the link between characteristics and Leave voting invites comparison with the statistical modelling used elsewhere in the literature. Becker et al. (2017) identify four broad hypotheses proposed as key drivers of the Leave result: EU exposure (trade, immigration and transfers), austerity and public service provision, demography and education, and economic characteristics (economic structure, wages and unemployment).[5] For local authority-level data their best subset selection procedure highlights education profiles, skill levels and measures of deprivation as important linear predictors of Leave voting. Alabrese et al. (2019) apply a similar variable selection approach to Becker et al. (2017), both individual and local authority-level data, identifying demographic and employment characteristics as the strongest predictors of the Leave vote, alongside significant geographical heterogeneity.

However, some, or all, of the many explanatory factors proposed as drivers of the referendum outcome may interact in a non-linear fashion. Studies premised on linear relationships or bivariate considerations may miss important elements of the story, or stories which appear only when a fuller set of interactions and non-linearities are captured. Statistically modelling such requires additional coefficients, costing degrees of freedom and is restricted by the inherent multicollinearity in spatially aggregated demographic data. In environments with limited observations introducing further covariates is impractical, imposing additional constraint on what may be empirically evidenced. For example, Zhang (2018) focuses only on local authority percentages with degrees, the unemployed and upper social class. Such modelling strategies enable ordinary least squares (OLS) analysis but leave questions around the role of the omitted characteristics. TDABM avoids compromise of this nature as the representation only shows how voting behaviour varies across the characteristic space, which may be highly multidimensional without loss of degrees of freedom, and no causality is assigned.

Exploration of interaction effects often reveals important nuances. Antonucci et al. (2017) find the negative overall correlation between education and the Leave vote is driven by the strong association between intermediate education levels and Leave; the relationship is non-linear. This association is stronger in conjunction with perceived economic decline. Rather than the 'left behind', they attribute much of the Brexit vote to the 'squeezed middle', voters identifying as working class while holding middle-class jobs. Becker et al. (2017) also emphasise the need to see 'whether salient factors reinforced each other', though robustly investigating multiple interactions is impossible in an OLS set-up. Liberini et al. (2019) find measureables such as age important predictors of anti-EU sentiment not only in themselves but also in conjunction with other characteristics. They also show that individuals' feelings about their circumstances mattered more in referendum voting than their actual circumstances. In recognition, we include subjective well-being in our analysis (cf. Alabrese et al., 2019, who use life satisfaction measures). These interactions further motivate the use of model-free TDABM.

Our study takes its cue from repeated observations in the empirical literature that interactions and non-linearities have much to say on the question of why Brexit happened. We explore how similarity among constituencies based on a particular combination of multiple characteristics (leading to what we could term 'group belonging') translates into patterns of Leave and Remain support in the 2016 Referendum vote. Formation of these higher level groupings is based on non-spatial characteristics and a priori, there is no reason to find geographically close constituencies in the same group[6].

Our TDABM mapping emphasises the role of contextual effects in mediating individual-specific characteristics and generating spatial variation in voting behaviour. Such local shaping happens through politically and socially driven channels (Johnston et al., 2004). First, shaping may be politically motivated, locally targeted campaigning can directly affect local voting, or individuals may observe issues with their local socio-economic environment and then use the political process to pick a platform best aligned to resolving those. Second, shaping may be socially driven, happening through direct interaction or via social norms determined by the general features or attitudes of neighbourhood residents which individuals then emulate. Conversely, variation in voting across localities could result from self-segregation into like-minded neighbourhoods, leaving place itself (and hence place-based policy) a limited role. A large literature attests to the importance of neighbourhood effects in UK voting behaviours and weighs against ecological fallacies.[7] Abreu and Öner (2020) extend the characteristic set using British Electoral Survey individual respondent-level data on voter attitudes to produce constituency level aggregations. Contextual effects dominate compositional effects in explaining referendum voting: political context (the strength of the UKIP vote and local turnout) and values-based cultural 'grievances' (being anti-immigration, same-sex-marriage and redistribution) have strong effects, while economic context effects are less distinguishable. In particular, regardless of individual skills, the propensity to vote Leave is lower in areas with an above-average proportion of skilled workers. Likewise living in a culturally conservative area increases Leave-voting. Where mobility is low, constituency-average outcomes registered on socio-economic criteria may reflect access to opportunities in that place. Whether the constituency unit straightforwardly reflects the aggregated features of individuals who (choose to) live there, or features of local communities formed through neighbourhood effects channels, is not possible to disentangle in our study. However, the evidence discussed in Abreu and Öner (2020) gives weight to a contextual interpretation. We interpret the clusters of constituencies we find as 'types' of communities sharing a similar local context (as defined by their joint characteristics in the dataset) relevant for understanding the referendum vote.

The constituency groupings revealed in our study are interesting for analysing the Leave and Remain campaigns; both their respective advantages before the referendum, and how far their performance may have successfully converted voters to their side. Evidence from political science suggests that campaign effects themselves depend on interactions with voters' socio-economic characteristics (Hobolt, 2007; De Vries et al., 2011). Goodwin et al. (2018) argue that high information asymmetry in the run up to the UK referendum created potential for significant campaign effects, magnified by the absence of clear partisan cuing on Brexit from the major political parties, with both Labour and Conservative parties split on the issue.[8] In an online survey experiment conducted in 2015, they examine the effectiveness of pro- and anti-EU frames. While receptivity to pro-EU frames is especially associated with certain characteristics (Labour support,[9] under 26s and those undecided about the referendum), they find a significant interaction effect with education, with pro-EU arguments strongest among those with lower levels of education, though interactions with other socio-economic characteristics are not checked. The paper concludes on the failure of the Remain campaign to frame arguments adequately to persuade voters to their side.

Also interesting in this context are the findings of Shaw et al. (2017), whose content analysis of nine television debates in the weeks preceding the referendum reveals core differences between the campaigns. They conclude that 'Leave focused on a more consistent and tightly focused set of campaign themes, provided more explanation of those themes, and focused more on their own core issues than Remain' (p. 1020). They also analyse tactics employed in the debates by both sides, noting the tactic of tapping into emotion, in particular.[10] While it is not possible to capture the response of voters to emotive campaign messages, their premise is that voters could be influenced by the content of debates, placing an emphasis on the campaigns themselves and their *ex ante* ability to shift political opinion on the question of Brexit. On the other hand, micro-econometric analysis of the result has tended to point to more fundamental drivers. Becker et al. (2017) emphasise the explanatory role of 'variables that seem hardly malleable in the short run by political choices (variables such as educational attainment, demography and industry structure)' (p. 605). The conclusion need not be that the campaigns were irrelevant to the outcome, however. The results of our TDABM analysis of socio-economic characteristics at constituency-level reveals clusters of 'similar' places, and our suggestion is that these socio-economic clusterings represent the kinds of local context relevant for understanding how political campaign arguments and tactics are received, as well as for understanding constituencies' deeper predispositions on the question.

This paper thus represents the first mapping of the joint distribution of constituency level characteristics, identifying homogeneity in Leave voting behaviours that favours consistent campaign messaging and a stronger sense of belonging. Noting that identical regression coefficients can be obtained from data with very different distributions (Anscombe, 1973), the inference on the joint distribution of constituency characteristics enhances understanding beyond any evidence provided by regression models in the extant literature.[11]

With results robust to indicator selection and spatial aggregation level, we argue the consideration of the topology of regional characteristics is an important part of understanding both the Brexit vote and any subsequent treatment of constituencies premised on the 'EU discontent' and 'left behind' narratives.

3. DATA

As a base for analysis we use parliamentary constituency data as compiled by Professor Pippa Norris for analysis in Thorsen et al. (2017).[12] Demographic data from the 2011 Census data are merged in Thorsen et al.'s set, permitting the analysis of the socio-demographic space upon which the Brexit vote played out. Because constituencies do not correlate directly with counting districts used in the referendum, Hanretty (2017) constructs estimates of the percentage of voters for Leave and Remain in each constituency. Robustness checks demonstrate that the model used to convert local authority-level voting data to the constituency level does not impact the inference derived from the TDABM analysis. We combine the 2019 General Election results with results for 2015 and 2017, allowing consideration of the EU referendum in the context of changing party affiliation.[13] The election in 2015 represents the last election before the referendum and indicates prior political leanings of constituencies.

The constituency characteristics we investigate are housing tenure and occupancy, motor vehicle access, NSSEC status, qualification levels and self-reported health. Our dataset contains information for 45 different categories within these seven questions reported from the 2011 Census.[14] In common with Becker et al. (2017), Hanretty (2017), Abreu and Öner (2020) and many others, the assumption is that constituencies' demographic make-up did not vary greatly between 2011 and 2016. While question-by-question analysis may be interesting, the 2016 Referendum results are the consequence of all factors in combination. TDABM is a big data approach and can readily extend to 45 axes without undue concern for degrees of freedom. However, categories for which many constituencies register very low proportions often create connections, as the balls encompass their full range long before the other characteristics.[15] One treatment, as applied in the machine-learning literature is to normalise all variables onto the range [0,1], but this then gives equal importance to all variables in the plot. Since all values in this paper are on the scale 0–100%, we do not rescale here. Instead, we have slightly reduced the number of axis variables by merging some categories where there is a rationale to do so. For instance, in any constituency a very small proportion of households own three or four cars. We consider the proportion of

households with two or more cars on the basis that this is the more salient information. We also merged categories which tend to be distributed in similar areas of the TDABM plot (e.g., NSSEC categories and self-reported health; see the supplemental data online for TDABM plots coloured by axis variable). Certain categories making up very small proportions for all constituencies are dropped, for example, the proportion of households for which all household residents are 65+. Age categories are included as separate axis variables.

The 27-axis variables employed in the analysis are listed in Table 1 along with descriptive statistics for the average proportion of each category in each constituency. These provide a flavour for the data and permit some basic testing of links with Brexit voting. Using two-sample *t*-tests we report the difference between the values of each characteristic in Remain and Leave areas, reporting a positive value in the 'difference' column whenever the average value amongst Leave is higher. Across the dataset, broad averages are consistent with correlations reported in existing literature on the Brexit vote: lower education levels, higher levels of deprivation, lower skilled occupation and poorer health are all associated with the Leave vote, supporting an exclusion story whereby those who felt socioeconomically 'left behind' drove the referendum result (e.g., Goodwin & Heath, 2016b; Mckenzie, 2016; Hobolt, 2016; Inglehart & Norris, 2016; Bromley-Davenport et al., 2018).

4. METHODOLOGY

TDA views data as a point cloud, with the position of each datum point within the cloud defined by its values on each of the axes that comprise the cloud. With only two dimensions, that point cloud is analogous to a scatter plot. Each variable in the dataset may become an axis, subject to the requirement that the realised values within the dataset are ordinal and have sufficient variation. In this paper the cloud comprises the 27 characteristics of constituencies outlined in Table 1. As the dimensionality increases, a visualisation tool becomes necessary to obtain the valuable inference that is offered by scatter plots. TDABM addresses this, providing an abstract representation of the point cloud that maintains full connection to the underlying dataset.[16]

First, the TDABM algorithm selects a point at random from the dataset and constructs a ball of radius ε around it. In two dimensions a ball is simply a circle, but TDABM operates in any number of dimensions. The initial selected datum point is the first landmark. Any other data points within the ball are considered to be covered by the ball. TDABM will then select a second landmark from the uncovered set, marking as covered any points within a ball of radius ε surrounding that landmark. Continuing to iterate point selection, the algorithm finishes when there are no uncovered points.

Relative positioning of the balls is obtained through the presence of points in the intersection of two balls. Where there is a non-empty intersection, an edge is drawn between the two landmarks. Density of the cloud is captured by resizing the representation of the ball to reflect the number of points it contains; larger balls signify more points within radius ε of the landmark. Summary statistics may be produced for each ball including the average value of the outcome of interest for the points within it. In the visualisation, balls are coloured according to a function of the points they contain. Primarily in this paper colouration is according to the average Hanretty (2017)-estimated Leave percentage. We also tell a voting story by colouring by outcomes of the general elections of 2015, 2017 and 2019. The resulting TDABM plot shows the joint distribution of the characteristics and, as demonstrated by Anscombe (1973) for the single-axis case, is independent of any modelling that links the independent characteristics to the dependent variable used as colouration.

Outcomes from the TDABM algorithm are dependent solely on the choice of the radius ε. Just as there is no single optimal scale at which to view a geographical map, so no hard rule exists about the optimal choice, but it is straightforward to iterate over radii to verify the robustness of conclusions. Outcomes are also dependent on the random landmark selection. However, as shown in subsequent sections, the broad inference of TDABM is consistent over multiple applications of the algorithm with different random seeds. Visualisations can thus be understood with confidence and bootstrapped confidence intervals constructed on any metric derived from the TDABM plot. In what follows ε = 23 is used; a demonstration of the strong robustness of the key messages in this paper to ε selection is provided in the supplemental data online.

5. RESULTS

5.1. BallMapper results

Figure 1 provides a TDABM graph with ε = 23, showing a large concentration of balls to the centre left, with three arms extending towards the bottom and right. Colouration is by average Brexit support in each ball. Each ball is a collection of constituencies with broadly similar characteristics from our combined category set. A join between two balls means that there is at least one constituency sitting in both balls. As this plot seeks to represent 27 dimensions in two-dimensional form, there is no direct interpretation of the horizontal or vertical direction. The TDABM graph allows us to see the shape of the data; to find out more about specific variables' behaviour we would colour by that variable. The supplemental data online provides plots coloured by each axis variable; plots in Figures 1 and 3 are coloured by non-axis variables which can be thought of loosely as 'outcome' variables.

Three points emerge immediately from Figure 1. First, the comparative concentration of constituencies within the upper left of the plot. Though this has no direct interpretation in terms of the values of the demographic indicators, it does inform that the constituencies here are very similar to each other in all the 23 considered dimensions. There is only one disconnected ball, informing that most

Table 1. Summary statistics and univariate tests.

Question	Variable	Mean	SD	Minimum	Maximum	Leave versus Remain			Correlation
						Leave	Remain	Difference	
2015 Vote (%)	Labour	32.35	16.50	4.51	81.30	31.65	33.58	−1.93	−0.06
	Conservatives	36.66	16.16	4.67	65.88	39.60	31.46	8.14***	0.24
	Liberal Democrats	7.82	8.36	0.75	51.49	6.96	9.34	−2.38**	−0.26
	Others	23.17	11.85	6.09	65.33	21.78	25.62	−3.83**	
Housing tenure	1: Owned	64.05	11.42	20.48	85.68	67.04	58.78	8.26***	0.44
	2: Social rental	17.99	7.80	4.59	50.63	16.73	20.21	−3.48***	−0.23
	3: Private rental	15.90	6.41	5.55	42.10	14.26	18.80	−4.54***	−0.47
	4: Other	2.06	0.65	0.82	7.93	1.97	2.21	−0.24***	−0.22
Household composition	1: Married	33.33	5.76	14.63	46.33	34.43	31.39	3.04***	0.35
	2: Cohabit	9.74	1.49	3.50	13.82	10.05	9.19	0.86***	0.25
	3: Other	45.91	7.35	32.26	71.85	43.82	49.61	−5.78***	−0.48
Car ownership	0 Cars	25.54	11.57	7.86	66.70	22.84	30.32	−7.49***	−0.40
	1 Cars	42.30	2.99	28.24	50.25	42.78	41.46	1.32***	0.31
	2+ Cars	32.16	10.9	4.38	57.96	34.38	28.22	6.17***	0.34
NSSEC	1: (see notes)	40.04	8.60	20.25	64.58	38.28	43.16	−4.88***	−0.39
	2:	19.99	2.76	10.59	26.88	20.99	18.22	2.77***	0.61
	3:	25.82	6.47	9.73	40.31	28.26	21.51	6.74***	0.65
	4:	5.42	2.87	1.62	22.59	5.27	5.69	−0.42	−0.05
Qualifications	1: None + L1	47.49	4.01	35.22	60.40	45.51	50.99	−5.48***	−0.78
	2: L2 + apprenticeship	33.62	2.13	26.81	38.38	34.45	32.15	2.30***	0.59
	3: L3 + L4	18.89	3.35	10.96	29.38	20.04	16.86	3.18***	0.56
Self-reported health	Very good	37.61	8.44	15.64	67.00	39.55	34.18	5.37***	0.47
	Good	18.61	3.38	8.00	23.80	20.43	15.38	5.04***	0.78
	Other	38.81	8.75	21.31	67.68	35.07	45.42	−10.36***	−0.74
Deprivation	0	42.14	6.98	22.21	59.71	41.39	43.47	−2.08***	−0.22
	1	32.56	1.75	28.14	38.38	32.66	32.37	0.30	0.09
	≥ 2	25.30	6.29	12.15	44.77	25.95	24.16	1.78**	0.21
Age	1: < 18	21.10	2.39	12.54	33.79	21.41	20.56	0.85***	0.25
	2: 18–24	9.30	3.59	5.73	32.68	8.50	10.70	−2.2***	−0.35

(*Continued*)

Table 1. Continued.

| Question | Variable | Mean | SD | Minimum | Maximum | Leave versus Remain | | | |
						Leave	Remain	Difference	Correlation
	3: 25–59	46.65	3.41	36.49	63.23	45.74	48.26	−2.53***	−0.48
	4: ≥ 60	22.95	5.34	7.78	40.56	24.35	20.47	3.88***	0.43

Note: Variables are organised by question; the total for each question is 100%. All variables from the 2011 Census except 2015 vote percentages are from Thorsen et al. (2017). Categories are combined from individual answers from census data. Owned housing tenure combines owned outright and mortgage categories; 'other' household composition includes living alone, lone parent, all-student households and all others. National Statistics Socio-Economic Classification (NSSEC) categories are (1) higher managerial + higher professional + small employer; (2) intermediate + lower supervisor; (3) semi-routine + routine; and (4) never worked + long-term unemployed. 'Correlation' is between the proportion of individuals responding to the 2011 Census as belonging to a category and Hanretty's (2017) estimated Leave percentage. Leave versus Remain divides constituencies on whether Hanretty's estimated leave percentage is greater (smaller) than 50%. Difference is augmented by the significance of a two-sample *t*-test for equality of means: *5%, **1% and *** 0.1%.

constituencies have at least one other to which they are quite similar. Second, the Leave vote, coloured on the yellow to red scale, is concentrated within a core part of the space to the left of the plot. Leave particularly covers the larger balls in the upper left. Finally, the Remain colouration, on the blue scale, sits away from the main mass of the plot and is more thinly spread. This tells us there is greater heterogeneity between Remain voting constituencies. We return to this important observation subsequently.

Table 2 provides average Remain percentages and numbers of constituencies for each of the balls in the diagram. Together with Figure 1, the recurring message is that there are more Remain-supporting balls than Leave: that is, there are more balls where the average Han-retty-estimated Leave percentage is lower than 50%. Furthermore, of the 11 balls containing more than 50 constituencies only one, ball 8, has an average estimated Leave percentage below 50%. Discounting ball 8, the average number of constituencies in a ball with less than 50% esti-mated Leave vote is just 9.5.[17] By contrast the average size of a ball with greater than 50% estimated Leave percentage is 68.6 constituencies. These statistics reinforce the mess-age of Leave homogeneity which emerge from viewing the TDABM plot.

In line with related evidence, aside from the concen-tration of Scottish and London-based constituencies discussed below, we find no strong regional pattern across England and Wales and so emphasise place at constituency level in what follows (Johnston et al., 2018).[18] TDABM output retains the data points in each ball for interrogation. Ball 32, to the bottom centre of the shape, contains Glasgow East, Glasgow North East and Glasgow South West, seats held by the pro-EU Scottish National Party (SNP) (Henderson et al., 2021). Glasgow South West provides a characteristic link into many other similar constituencies from other industrial cities in ball 37 which features areas such as Bootle, Walton and West Derby in Liverpool, Blackley and Broughton in Manchester, and Birmingham Erdington and Nottingham North. We may view this arm as being the more deprived areas of major cities, with low qualifications, few higher professionals and a higher propensity to rate health lower (a type of location linked to Leave voting in Goodwin & Heath, 2016b; and Lee et al., 2018). We refer to this string of balls as 'arm A' of the plot.

The string of Remain balls running through 38, 20, 18 and 31 contains more cities. This string is labelled as arm B. Ball 38 has Nottingham East and Nottingham South, Sheffield Central and Newcastle upon Tyne East. These are Labour seats that were held in 2019. Relative to ball 32, ball 38 mainly differs on the high proportion of highly qualified individuals. Deprivation in ball 38 is also much lower than in ball 32. Ball 20 contains Nottingham East as a bridge. It has a higher level of deprivation than ball 38, but maintains the high proportion of residents with post-compulsory education. Moving along the arm into 18 and 31, the NSSEC levels of the jobs move lower
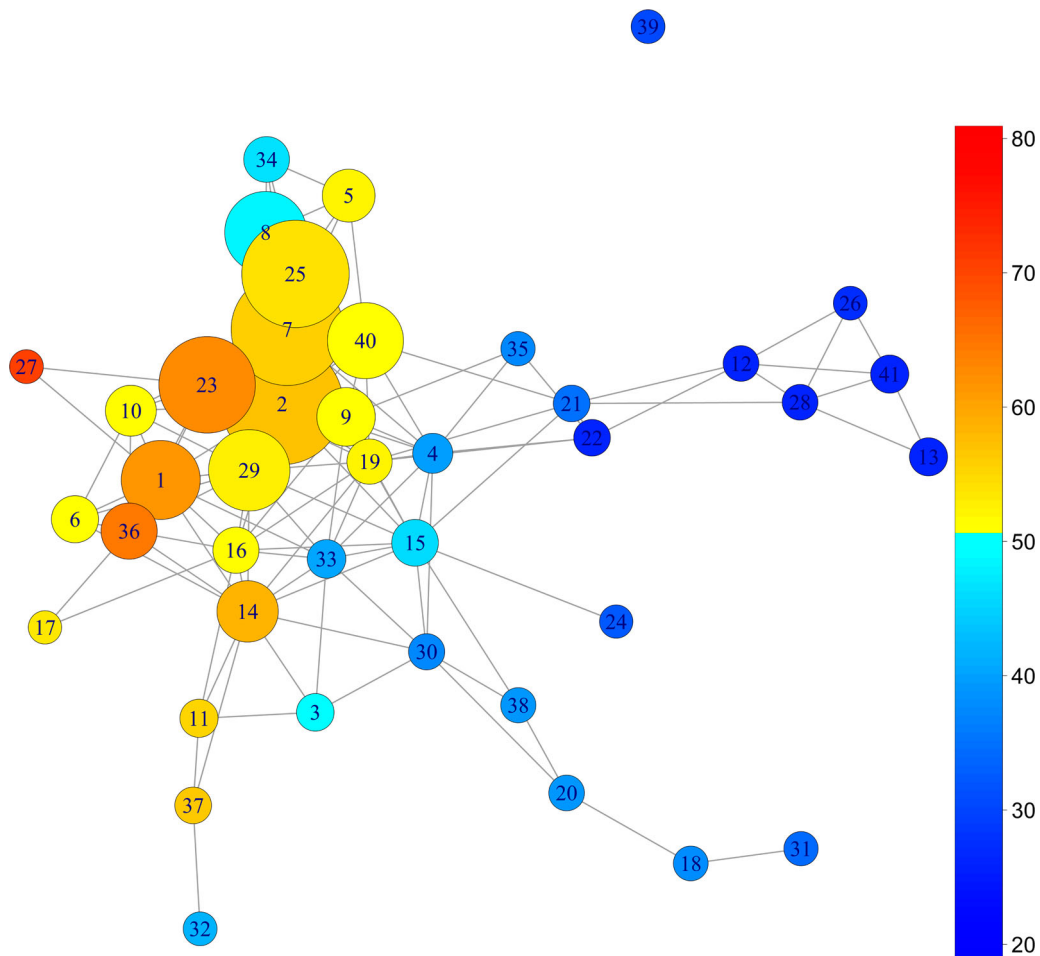
Figure 1. Leave vote percentages ($\varepsilon = 23$).
Note: Topological Data Analysis BallMapper (TDABM) diagram representing a 27 dimensional space of constituency characteristics using a reduced set of Census 2011 variable categories. Details of the combination of categories are provided in the data section. Colouration is by Hanretty's (2017) estimated Leave percentages with the 50% cut-off being towards the upper end of the shading. Axes are constructed from the combination of categories within questions.
Note: Readers of the print article can view the figures in colour online at https://doi.org/10.1080/00343404.2023.2204123
Source: Data are from Thorsen et al. (2017). Construction of diagram using R package BallMapper (Dłotko, 2019b).

and the levels of deprivation rise. Ball 18 picks up Manchester Central, Leeds Central, Tottenham and West Ham. Ball 31 is then Glasgow Central, Leeds Central, Manchester Central and Liverpool Riverside. These are very different communities to those of ball 32, with ball 31 having more young adults, higher incidences of private rental, higher qualification levels and more residents at either intermediate or lower supervisor on the NSSEC classification. This arm captures many of the communities in and around universities linked to Remain support in local authority-level analyses (Johnston et al., 2018; Zhang, 2018). Balls are very similar in having low car ownership, low self-reported health and higher proportions in the household composition group that combines those living alone, lone parents and households of students.[19]

The strong Remain arm, heading out from balls 4 and 21, features constituencies such as Bristol West, Manchester Withington, Hove and Brighton Pavilion in ball 22.[20] This final arm is labelled as arm C. Ball 13 is entirely London boroughs and includes Islington North, Hackney,

Bethnal Green and Bow, and Hammersmith. Ball 26 contains the Cities of London and Westminster, Kensington, and Hampstead and Kilburn, more affluent areas of London that contrast strongly with the boroughs of ball 13. Kensington forms the overlap with ball 41. Also in 41 are Hackney North, Vauxhall, Lewisham and Deptford, Hammersmith and Islington North, all Labour seats within London. The differences between balls centre on deprivation, qualifications and the extent of social rent. Moving to the right of this arm is increasing deprivation, greater prevalence of social rent and reduced car ownership. However, starting from 21 the balls along this arm feature the highest mean values of level 4 qualifications in the plot.

Finally, we note ball 39 as an outlier Remain ball containing Richmond Park, Twickenham and Wimbledon. These are all suburbs of West London with high levels of home ownership, high qualifications and low deprivation. The age distribution is more skewed towards older residents and occupations tend to be from higher NSSEC groups.
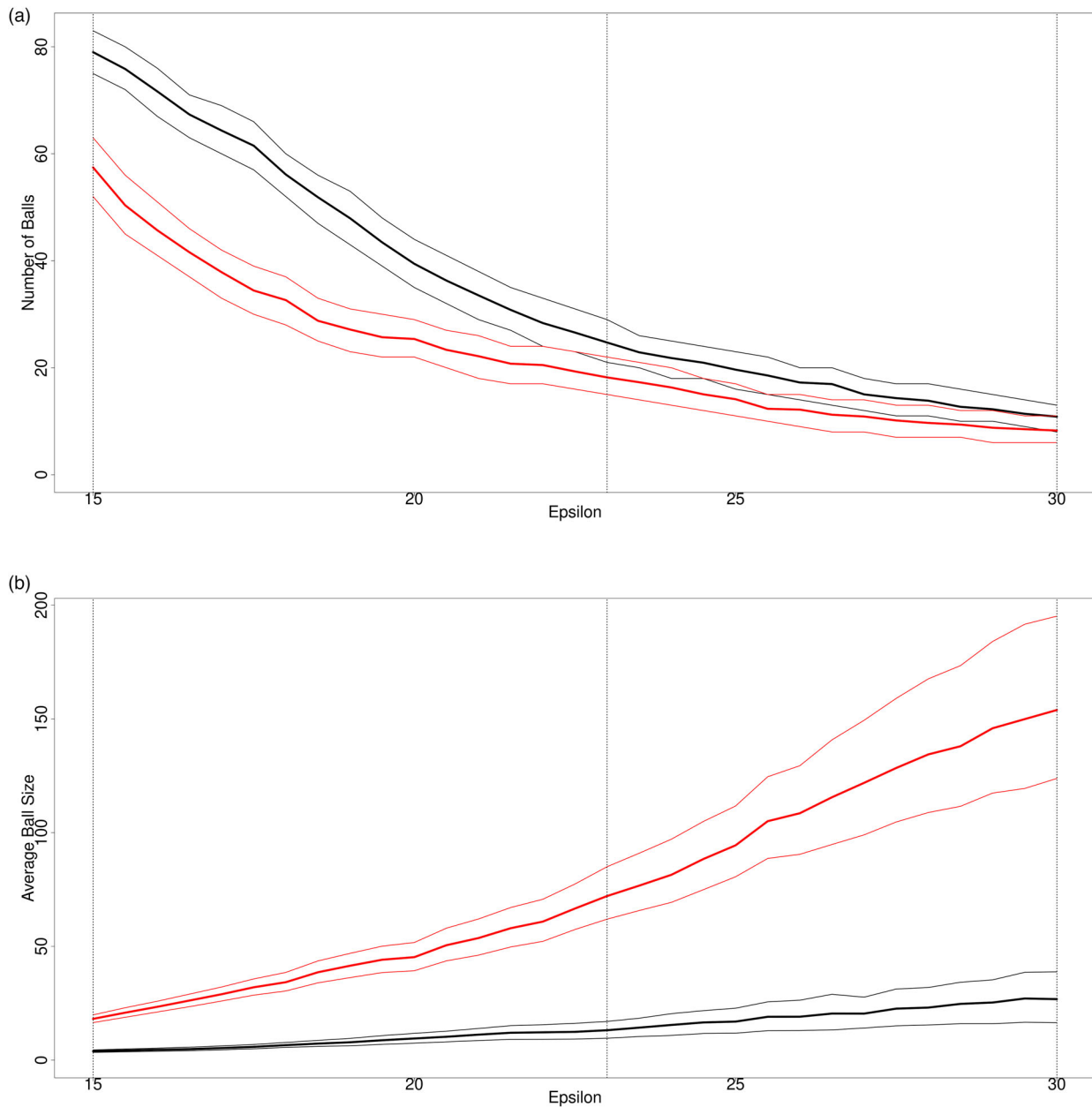
Figure 2. Brexit constituency distribution robustness: (a) high numbers of balls at low radius and low numbers at high radius; and (b) average number of constituencies in a Remain ball and a Leave ball.

Note: Figures plot the average ball size and number of balls that satisfy the condition that the average Leave percentage for the ball is less than 50% and greater than 50%. These averages are based on 10,000 repetitions of the topological data analysis BallMapper algorithm at radii $\varepsilon \in [10, 30]$ in increments of 1. Lines relating to Remain balls are plotted in black, whilst lines relating to Leave balls are plotted in red. Data are 95% confidence intervals from the 10,000 estimates plotted as thinner lines in the respective colours. For (a), the numbers of leave and remain balls get smaller, and closer together, as the radius increases; and for (b), the numbers get higher as the radius increases.

The three Remain arms are all very different, hence we do not see any connectivity between them. There are the more deprived constituencies of Glasgow in arm A, the diverse regional city centres in arm B and the London boroughs in arm C. There is a converse similarity about the Leave voting areas, the oranges and yellows on Figure 1. Towards the top of the main body of balls are balls 5 and 25 with Hanretty-estimated Leave percentages of 52% and 54%, respectively. In the centre left are 23, 1 and 36 that are deeper orange in colouration and have

Hanretty-estimated Leave percentages of 62%, 62% and 64%, respectively. These are very strongly Leave-supporting balls. Finally there is the yellow colouration stretching down into arm A, balls 11 and 37 having Hanretty-estimated Leave percentages of 55% and 57%, respectively.

Balls 25 and 5 are predominantly rural with very similar Hanretty-estimated Leave votes, containing constituencies such as the Derbyshire Dales, Central Devon, South Suffolk and West Worcestershire, also towns such as Aylesbury, Newark, Shipley and Stratford-on-Avon, all
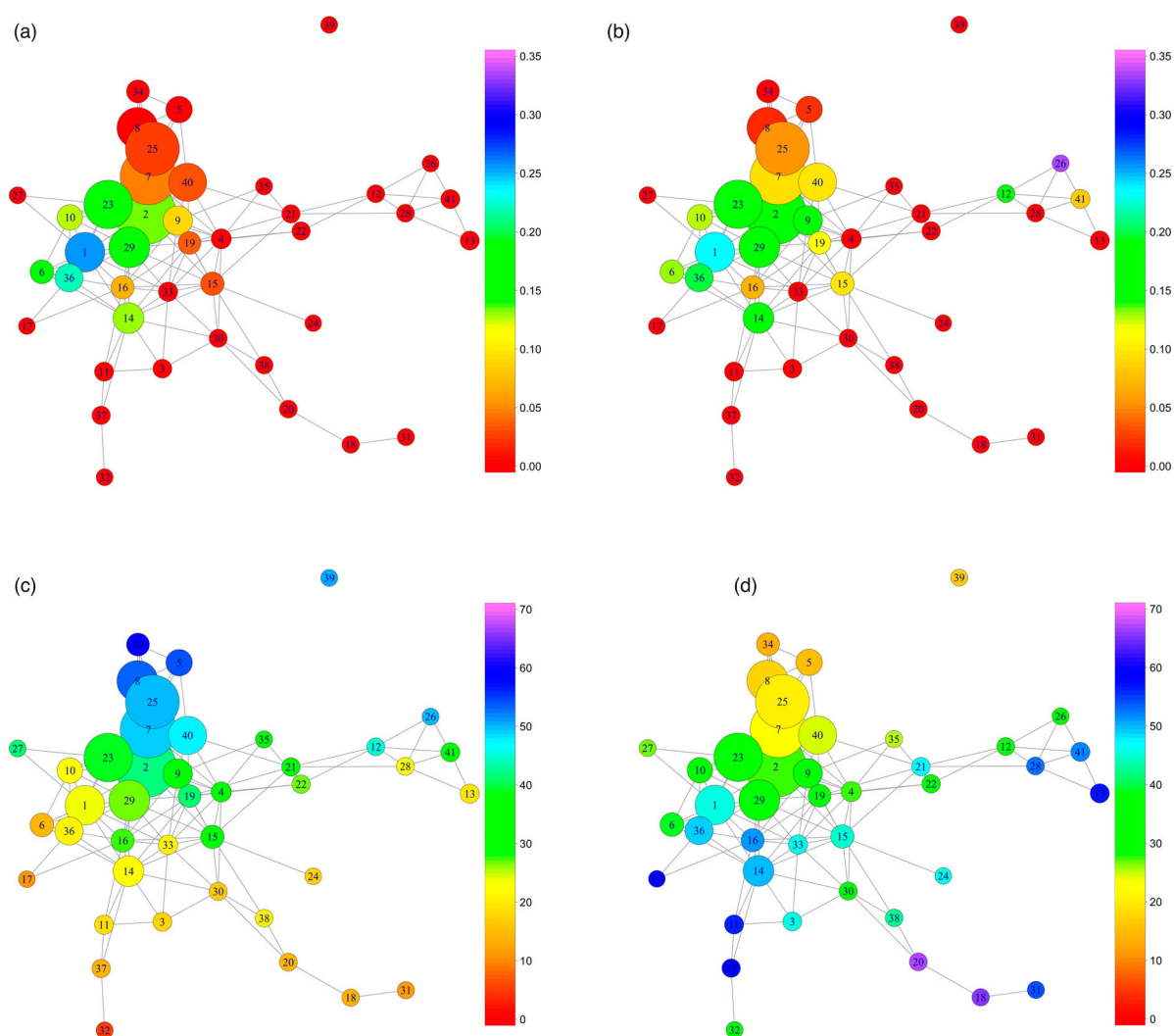
Figure 3. Conservative Party gains from the Labour Party at December 2019 ($\varepsilon = 23$).
Note: Topological Data Analysis BallMapper (TDABM) diagram constructed using *Bal*lMapper (Dłotko, 2019b). Panels (a, b) are coloured by proportion of constituencies in each ball won by the Conservatives from Labour in the December 2019 General Election relative to the 2015 and 2017 General Elections, respectively. Panels (c, d) are coloured by the 2015 vote shares for Conservatives and Labour, respectively.
Sources: Data are from Thorsen et al. (2017) and www.gov.uk.

with marginal votes to Leave. As may be understood with the overlap to ball 8 there are constituencies whose vote was marginally in favour of Remain that also sit in this ball, such as North Somerset, Monmouth and Horsham. In this area of the plot home ownership is high, having two or more cars is common and the highest NSSEC occupations are found; qualifications are high and self-reported health is very good. Models have aligned many of these characteristics with Remain, but as we see the overall combination leans to Leave. Constituencies in ball 8 with Hanretty (2017) estimated Leave percentages above 50% may be understood as the 'squeezed middle' discussed by Antonucci et al. (2017). Regions of the TDABM plot like this highlight the importance of interactions within the data. Visualisation facilitates a reappraisal of the understood relationships from additively separable regressions.

Balls 1 and 36 are where some of the strongest average Leave vote is found. Here the constituencies are

predominantly urban, with many linked to industrial decline. Leave voting in such places has been linked to their particular exposure to disruption from trade liberalisation and a subsequent narrative formed around economic grievance (Los et al., 2017; McCann & Ortega-Argilés, 2021) These balls contain Blackburn, Burnley and Bradford of the traditional textile towns; also former coalmining areas such as the constituencies of Merthyr Tydfil, Normanton, Pontefract and Castleford, Rhondda and Bolsover; and the former steel areas of Scunthorpe, Redcar, Stocksbridge and Rotherham. These are constituencies where health is poorer, qualifications lower and deprivation is high. However, there are similarities with balls 5 and 25 too: high home ownership, marriage and similar prevalence of one car households and upper-middle NSSEC class occupations.

Moving down into arm A and towards ball 32 we find balls 11 and 37, containing deprived suburbs and a mixture

Table 2. Leave vote percentages summary ($\varepsilon = 23$).

Ball	Size	Leave	Ball	Size	Leave	Ball	Size	Leave	Ball	Size	Leave
1	98	61.57	12	7	26.34	23	135	62.29	34	28	46.81
2	188	57.00	13	12	26.72	24	2	32.27	35	4	37.63
3	11	49.31	14	61	58.16	25	158	54.21	36	50	64.33
4	16	39.81	15	30	45.74	26	3	27.66	37	9	56.51
5	43	52.20	16	29	51.48	27	3	70.49	38	6	39.29
6	31	51.11	17	2	53.32	28	7	26.70	39	3	30.46
7	168	55.68	18	5	37.94	29	103	52.87	40	92	51.03
8	104	48.45	19	27	51.99	30	8	37.67	41	12	26.54
9	55	50.94	20	7	39.35	31	4	34.59			
10	39	51.81	21	9	35.27	32	3	41.79			
11	12	54.93	22	9	26.56	33	13	40.57			

Note: Ball numbers are related to the topological data analysis BallMapper plot of our reduced category dataset with $\varepsilon = 23$ plotted in Figure 1. Size is the number of constituencies contained within the ball. Leave is the average Hanretty (2017)-estimated Leave percentage for the constituencies contained within the ball. These Leave values correspond to the colouring of the balls in Figure 1.

of voting behaviours. Ball 11 contains Gateshead, Leeds East and Birmingham Erdington, all with Hanretty-estimated Leave percentages above 55%, alongside Edmonton and Newcastle-upon-Tyne Central with Leave percentages below 50%. In ball 37 are Bootle, Middlesbrough and Nottingham North, all with very high Leave percentages. There is then Glasgow South West which serves as the link into the blue coloured ball 32. These constituencies have lower home ownership, the dominant category being social rent. There are lower levels of marriage and more households without access to a car. Health and qualifications are lower and deprivation higher.

5.2. Remain heterogeneity, leave concentration

Primary inference from Figure 1 is that Leave support is far more concentrated within the space than Remain. The scale on the right reporting Leave percentages as estimated by Hanretty (2017) places 50% between the light and dark blues; all Leave constituencies are in the centre of the big shape in the left part of the plot. It is also immediate that the biggest balls correspond to those voting to leave the EU, while those wishing to remain are more spread out across smaller balls. The strongest Remain constituencies are found on arm C, extending out to the right, though each arm goes to Leave percentages of less than 40%. There are more marginal Remain areas to the top and left of the main connected shape. That Brexit-favouring constituencies are more jointly similar on these axes than others comes through strongly in the plot.

We investigate the robustness of Remain heterogeneity versus Leave concentration using the TDABM algorithm. By iterating the TDABM algorithm 10,000 times over radii in the range $\varepsilon \in [10, 30]$. we can understand more about the nature of Leave and Remain balls. From each iteration we collect the average size of balls that have a colouration value less than 50%. We also collect the number of balls that average a Remain vote from each TDABM graph. Figure 2 shows the results.

Figure 2 has two panels. In each panel the black lines relate to the Remain balls, whilst the red lines relate to those balls with an average Hanretty (2017) estimated Leave percentage above 50%. The left panel reports that the number of Remain supporting balls is consistently higher than the number of Leave supporting balls. Likewise the size of the Remain supporting balls is much smaller than the average size of the Leave supporting balls. Confidence intervals from the 10,000 repetitions confirm these results to be significant. Consequently our illustration in Figure 1 is no exception in showing the Remain concentration.

Turning this towards an understanding of the Remain campaign's failure, we focus on voters in marginal areas. Marginal areas are coloured in light blue in Figure 1: balls 8 and 34 form one marginal pair, while a line of others cuts through the plot from ball 3 to ball 35 through balls 33 and 4. The prominent Remain areas are then to the right of this line. We could ask: What campaign message could have converted constituencies in those light yellow-coloured balls, like 6 and 40, to higher Remain support? Our contention is that if campaign messages are more successful when targeted to the community's joint characteristics, a message targeting ball 6 would not simultaneously convert those in ball 40 (while retaining core supporters in ball 30). What might be effective in mobilising votes in constituencies in one part of the space may not be effective elsewhere. This diversity necessitates different messages and opens potential for conflicting signals that diminish impact. The conclusions of Shaw et al. (2017) regarding the relative 'incoherence' of the Remain campaign at the national level are then less surprising.

Analysis of the TDABM results pointed to the clear heterogeneity of the Remain voting clusters. The variation from Glasgow suburbs to the centres of the major cities and to the boroughs of London could not be starker. Not only are the geographical and political differences clear, but also the overall difference across joint characteristics is large. As the outcome of the referendum is based

on the total nationwide vote, seeking extra support within any constituency has merit.

Balls indicate where marginal constituencies sit in the characteristic space. As an example, consider ball 19, itself yellow but connected to blue-coloured balls 4, 15 and 33. Ball 19 contains smaller urban areas estimated to have voted Remain, such as Cheltenham, Chester, Exeter and Hove, but also other similar rural conurbations estimated to vote Leave such as Colchester, Lincoln, Poole and Worcester. Moving up from ball 19 into balls 40 and 7 deprivation falls, but moving left into balls 29 and 2 the proportion of households suffering two or more of the deprivation indicators rises. Ball 19 does not contain so-called 'red wall' seats where the Leave campaign had strong appeal (Harris & Charlton, 2016; Antonucci et al., 2017; Los et al., 2017), rather this is a set of constituencies where Remain messages had chance to resonate. The plot there-fore serves as a useful post-campaign evaluation tool.

6. FURTHER ANALYSIS

Our results show the contrasting nature of Leave and Remain support, the former concentrated in a small area of the socio-demographic characteristic space while the latter is highly spread; in other words, when all inter-actions among variables are taken into account, Leave-supporting constituencies are more alike than Remain constituencies. Colouring the plot by the 2015, 2017 and 2019 election results, we now use TDABM to illustrate how the changing political landscape plays out on our characteristic space. Confirming the robustness of the con-centration of Leave voting regions compared with the het-erogeneity of Remain, we consider the full set of constituency data, analysing a dataset matching that used by Hanretty (2017) to compute the Brexit vote at the con-stituency level, consider the case incorporating cultural variables investigated by Abreu and Öner (2020), and replicate our main analysis using local authority level data and the actual Brexit vote counts from the 2016 Referendum.[21] Each discussion demonstrates further the value of visualisations from the TDABM algorithm.

6.1. Political parties and Brexit

The impasse in parliamentary proceedings during exit deal negotiations ultimately led to a third UK general election within five years. Though Brexit's disruptive effect on British politics was still unclear in 2017 (Johnston et al., 2018), Figure 3 helps to visualise how exactly the 2019 General Election panned out, and how its results link back to the Brexit question. Figure 3(a, b) are coloured by the proportion of formerly Labour constituencies won by the Conservatives in December 2019. Employing the same axes as Figure 1 facilitates rapid comparison of elec-tion voting patterns and Brexit voting patterns. To add reference we also colour by the 2015 election results in Figure 3(c, d).[22]

These plots tell a clear story, showing how longstand-ing political allegiances were disturbed by the referendum (Ashcroft, 2019; Cooper & Cooper, 2021). In terms of

campaign emphasis, the Conservatives' message was a simple 'Get Brexit Done', while Labour laid out a spend-ing programme directed at remodelling society (*The Guar-dian*, 2019a). Figure 3(a, b) show Labour losses to the Conservatives located in areas of the plot where Brexit support was strongest in Figure 1. Ball 1 contains the post-industrial areas, particularly in Northern England and South Wales. Balls 2, 23 and 29 also have about 20% of constituencies gained by the Conservatives shown in Figure 3(c), having had average Labour votes of more than 30% in 2015. These balls include Sedgefield, the former seat of Labour Prime Minister Tony Blair which fell to the Conservatives in 2019. Conservative gains versus both 2015 and 2017 are then concentrated in this part of the shape, not to the centre or right where the Leave vote was weaker, reiterating the centrality of the Brexit question to subsequent election outcomes, and indicating Conservative Party repositioning in response to the political shock of the EU referendum (Hayton, 2022).

Comparing Figure 3(a, b) with Figure 3(c) reveals that in balls 7 and 2 the Conservatives added constituencies socio-economically similar to those already held in 2015, while the bigger proportions in balls 23, 36 and 1 show a Conser-vative swing in constituency clusters with working-class characteristics traditionally associated with Labour (Figure 3d), the phenomenon sometimes described as the collapse of the 'red wall' (Cutts et al., 2020). These balls con-tain traditionally Labour-voting constituencies in the north of England, North East Wales and the Midlands and the plots emphasise the centrality of Brexit to some of the Labour Party's long-time faithful.[23] Commentary at the time pointed to the 'increasingly unstable alliance of Labour's left and centre, its remain and leave electorates, and its middle-class and working-class bases' (*The Guardian*, 2019b), and Figure 3(d) confirms that Labour party support in 2015 was much more spread out across the BM plot than Conser-vative support in Figure 3(c). Some strong Labour support in 2015 is found in the two strongly Remain arms of the plot stretching to the right (arms B and C), backing up the general link between Labour party affiliation and propensity to vote Remain reported in the literature (Alabrese et al., 2019; Goodwin et al., 2017). However, the TDABM analysis pro-vided here highlights visually the types of constituencies deviating from this general correlation.

Figure 3(a, b) are generally similar, though Conserva-tive gains relative to 2017 are stronger, reflecting growing popular frustration with parliamentary gridlock over Brexit and loss of confidence in the Labour leadership (Ashcroft, 2019; Cooper & Cooper, 2021). Most constituencies moving from Labour in 2017 to Conservative in 2019 were also Labour in 2015, though a good counter-example comes at the very right of Figure 3(b) in ball 26. This con-tains Kensington, a seat Labour had taken from the Con-servatives at the 2017 General Election. In Figure 3(a) it is coloured red, as Kensington is not a gain versus the 2015 result.

A strong link from Brexit to the incumbent govern-ment's policy around Levelling Up has been noted
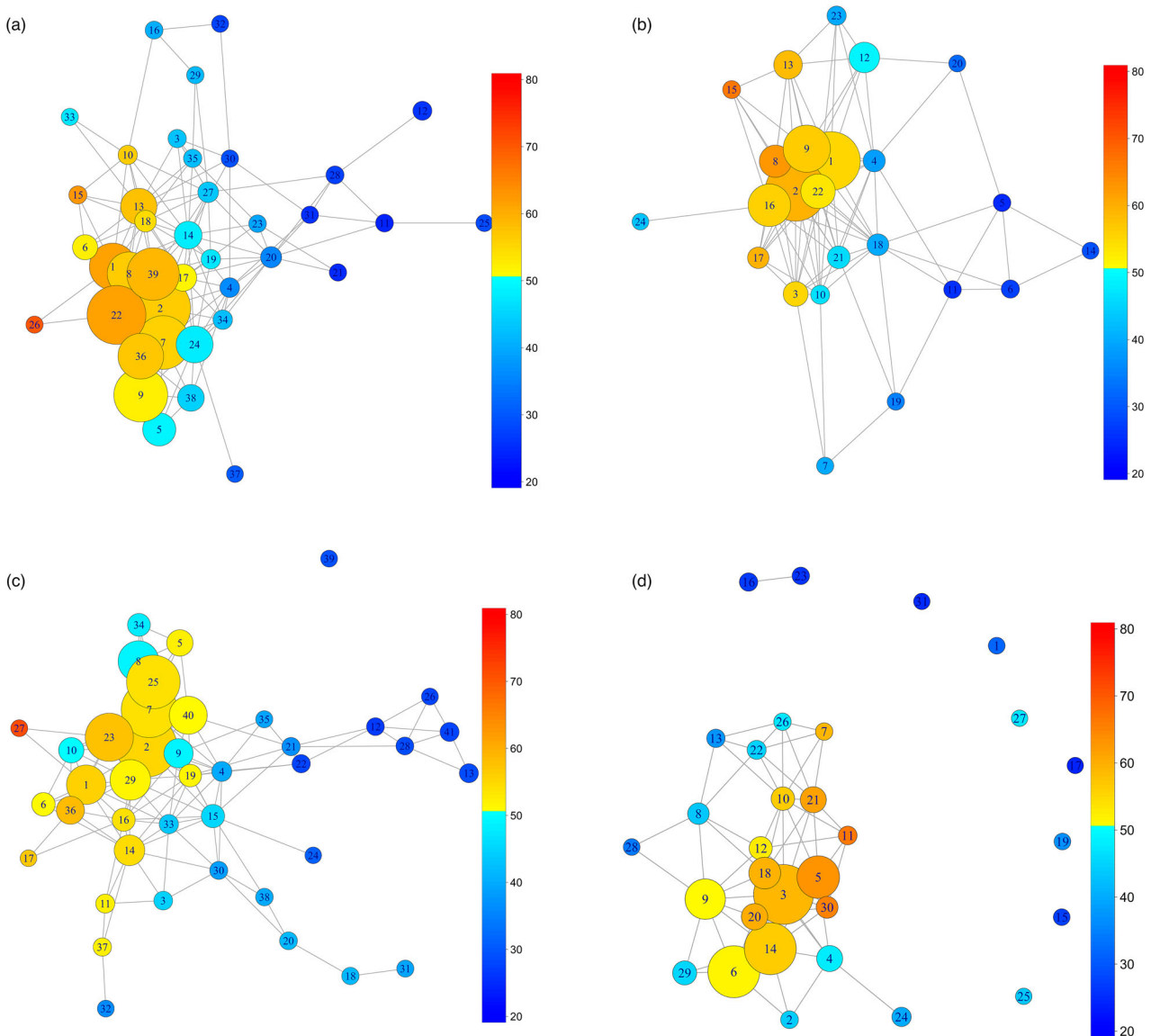
Figure 4. Alternative datasets.
Note: Topological Data Analysis BallMapper (TDABM) diagram constructed using *Ball*Mapper (Dłotko, 2019b). Panel (a) uses all variables from the Pippa Norris dataset as axes. Panel (b) incorporates the five cultural variables constructed by Hanretty and Vivyan (2015) in addition to the variables used in the main result. Construction of panel (c) is identical to the main results. Panel (d) uses the same variable construction as the main results, but with aggregation at the local authority level. Colouration of panels (a) and (b) is by Hanretty's (2017) estimated leave percentages. Colouration variable of panel (c) is according to estimates created by Hanretty and Vivyan (2015). Panel (d) is coloured by the actual observed Leave vote, which is available at the local authority level. In all cases with the 50% cut-off represented by the transition from lighter to darker colours.
Sources: Data are from Thorsen et al. (2017), www.gov.uk, Hanretty and Vivyan (2015) and NOMIS.

elsewhere; indeed, arguably, 'the 'Levelling Up' narrative was forged entirely out of the whole Brexit process and experience' (McCann & Ortega-Argilés, 2021, p. 521). Turning from political narratives to policy, a regional consequence of Brexit has been the withdrawal of EU Cohesion Policy funds. Analysis of funding allocation so far via the UK government's £3.6billion Towns Fund announced in 2019 is interesting in light of our discussion of Figure 3. The potential for funding to be allocated for political advantage through the Towns Fund is certainly there, owing to a competitive process judged against 'amorphous' selection criteria (McCann &

Ortega-Argilés, 2021; Wincott, 2021) and which has attracted recent criticism from the Committee of Public Accounts (House of Commons, 2022). Logistical analysis of town selection indeed reveals a strong link between being in a Conservative-held seat and the probability of obtaining funds, stronger still if that seat is a Conservative marginal (Hanretty, 2021). The Getting Building Fund and Community Renewal Fund show similar per capita funding boosts in Conservative-held areas after controlling for deprivation levels, with the highest premium for seats won in or after 2019 (van der Merwe & Goodier, 2022).

6.2. Alternative datasets

Figure 4 overlays voting patterns in the 2016 Referendum on four different datasets, varying the sets of characteristics and the level of spatial aggregation.[24] In all cases the colouration is set such that the orange colour scale applies at and above 51% voting for Leave. Blue colour scales apply at and below a 50% Leave vote.

Figure 4(a) uses the full set of characteristics that inform the combined categories of the main results. There are many similarities with Figure 1. The core of the shape features strong oranges – the high Leave vote that determined the result – while majority-Remain constituency groupings sit to the right, the blue balls which are smaller and more numerous (as before). While there are more interconnections between the Remain balls, the plot again displays a dispersed periphery. Indeed, there is a strong likeness between the shape of the two plots.[25]

Figure 4(b, c) incorporate additional data from Hanretty and Vivyan (2015). The plot in Figure 4(b) includes culture variables constructed from the 2014 British Electoral Study as additional axis variables.[26] These variables capture constituency attitudes to the redistribution of income, the impact of immigration on local culture, same-sex marriage, disillusion with the European Union and a scale of economic views from left to right. Although the distinct arms of the main results disappear, the heterogeneity of Remain and concentration of Leave is still evident. Figure 4(c) uses the same variables as the main results, but is instead coloured by estimated Leave proportions informed by individual-level survey data (Hanretty & Vivyan, 2015). Aside from balls 9 and 10 (now blue), the colouration aligns with Figure 1.

Figure 4(d) is constructed from local authority census data and coloured by recorded local authority Leave percentages. Again, this shows homogeneity of the Leave areas relative to the Remain areas, and demonstrates the independence of this result from the use of Hanretty's (2017) estimates. Figure 4(d) shows a series of outliers, all of which are Remain voting.

Changing the explanatory variables, incorporating political attributes, using pre-referendum voting intentions from the British Electoral Study, and using local authority level data with actual vote counts, all confirm the messages from the main results section. Namely, the alternative datasets show the Leave vote is concentrated in a denser region of the regional characteristic space, while strong heterogeneity exists in the Remain voting regions. These results apply whether aggregating at constituency or local authority level.

7. DISCUSSION

Though what Becker et al. (2017) acknowledge to be 'very simple empirical models' can explain a significant amount of the variation in the Leave vote share across local authorities, such models omit non-linearities recognised to be key and may obscure regional heterogeneities as well. Further, the models say nothing of the distribution of the characteristics upon which the coefficient estimates are based (Anscombe, 1973). Here we have instead taken a TDABM approach to investigate the grouping of constituency-level observations in a multidimensional space of socio-economic covariates, showing how Leave and Remain support varies across that map. Primary emphasis with this method is not on individual covariates, or trying to posit a linear relationship between them and the Leave vote (on which there are plenty of existing contributions), but on whether constituencies share things 'in common' in terms of those covariates taken together. Once commonalities are revealed, the researcher can dig down into the groupings to examine joint covariate behaviour in different parts of the space. This is a novel way of approaching the referendum data which accounts for multiple variable interactions. The richness of information in the TDABM plot offers empirical inference beyond that in the extant literature.

Constituency-level averages taken together can represent the multidimensional local socio-economic contexts in which voting decisions are made and as such may be highly indicative. Abreu and Öner (2020) have demonstrated that constituency context had large effects on 2016 Referendum voting beyond individual composition. Using constituency-level data, this paper has demonstrated that Leave-voting constituencies tend to share more commonalities than Remain-voting. That is not to say that all Leave-voting constituencies are alike – there is heterogeneity among them – but balls identified in Figure 1 as majority Leave-voting are larger, less numerous and more interconnected (indicating members common to more than one ball) than balls containing high proportions of Remain-voting constituencies. Those are small, more numerous and more spread out in the space. There is strong Leave homogeneity and Remain heterogeneity.

The TDABM plot established two distinct strings of strongly Remain-leaning groups: one linking diverse city centres, many with large university student populations, contrasting with the other string of London boroughs. Interestingly, further analysis shows how the combinations of shared characteristics change as we move along these strings from the centre towards the right; qualifications and NSSEC classifications fall and deprivation rises. A final string of balls highlights the Remain-voting constituencies of Glasgow in an outlying ball, linked on socio-economic characteristics to other deprived pro-Brexit constituencies of major cities. Meanwhile at the opposite end of the shape lie two groupings of affluent rural constituencies, sharing many characteristics with Leave-voting constituencies but nevertheless supporting Remain.

While we do not address causal mechanisms driving the Leave vote, considering the data from a different angle can suggest new avenues for modelling those channels. Our results suggest that group (ball) membership may be directly relevant for local propensity to vote Leave or Remain, alternatively to local openness to certain campaign messages or tactics on the Brexit question. This latter possibility casts the referendum campaigns as a channel linking socio-economic factors to referendum

voting behaviour; socio-economic factors in combination create a context in which particular types of political messaging are targeted or received.

Figure 1 reveals where the Remain campaign did not sufficiently convert opinion. Our results point not so much to a failure by Remain as to the comparative simplicity of the task faced by the Leave campaign, catering to relatively similar groups of constituencies while Remain-voting constituencies were highly diverse. To convert more marginal constituencies without alienating core Remain supporters would have required a more differentiated campaign; indeed, this may explain the relative incoherence of the national-level Remain campaign noted by Shaw et al. (2017). The TDABM plots recommend more active grass-roots campaign strategies. A glaring difference between Glasgow constituencies and adjacent Leave-voting groups, with whom they share so much, is the strongly coherent local messaging received from their SNP representatives, contrasting with the weak cues from main parties elsewhere. To this extent, viewing political geography against the economic geography context in the TDABM plot is illuminating.

Many critiques of data-driven approaches abide, and variable choice is clearly of great importance. Axis variables are selected here based on existing literature and the available data within the readily accessible dataset of Thorsen et al. (2017). However, the strength of the TDABM algorithm comes from the ability to deal in multiple dimensions. To that end the presentation here can be readily extended and an analysis of any ordinal constituency characteristic incorporated. Next logical steps would see the approach applied to multilevel datasets in order to investigate relative strength of compositional and contextual effects explicitly.

The study also provides further support for Brexit as an instigator of significant party change. Figure 3 illustrates the role of the EU referendum in redrawing the UK political map, with political parties repositioning more or less successfully in 2019 against changing patterns of allegiance among the electorate (Cooper & Cooper, 2021). The issue of Brexit remains high on the political agenda (Axe-Browne & Hansen, 2021), deciding as it did the Conservative majority in 2019. Voter sentiment on the question and narratives surrounding it continue to shape current government policy, in particular its Levelling Up agenda designed, in Boris Johnson's words, to answer 'the plea of the forgotten people, and the left behind towns' (Johnson, 2018, quoted in Wincott, 2021; see also McCann & Ortega-Argilés, 2021). It remains to be seen whether the coalitions built by the Conservatives in 2019 will hold into the next election, though funding allocation appears to have this objective firmly in view (Hanretty, 2021). Whether the centralised distribution of Levelling Up policy funds aligns with the recommendations of place-based policy theory, with its fundamental emphasis on socio-economic objectives, local capabilities and appropriate regional governance structures, continues to be challenged (Barca et al., 2012; Billing et al., 2021).

8. CONCLUSIONS

The relevance of local characteristics at various levels of regional aggregation has been widely discussed in the context of the decision of the UK to Leave the EU in 2016. This paper presents the joint distribution of constituency characteristics to map the Leave vote for the first time. Through our mapping of the socio-demographic space we illustrate the homogeneity of Leave voting constituencies versus the heterogeneity of Remain. We graph the geographies of discontent (Dijkstra et al., 2020; McCann, 2020) and the notion of 'left behind' regions (Goodwin & Heath, 2016a; Heath & Goodwin, 2017), defined on the combination of multiple socio-demographic indicators. Our results are robust to changes in the set of characteristics, aggregation of voting at the constituency level and to analysis of local authority data. Our approach is motivated by the recognition of interaction effects within the regional science literature (Antonucci et al., 2017; Goodwin & Heath, 2016b; Gordon, 2018), and the difficulties hitherto found for capturing interactions in a comparatively small dataset with many potential explanatory factors. Our contributions are obtained through an intuitive new data science methodology which is model-free and brings out the value of interactions among regional characteristics in shaping voting outcomes.

Within the sets of similar constituencies identified, heterogeneities remain. Some may be explained by higher level geographic aggregations, such as the tendency of Scottish constituencies to vote Remain (Clarke & Whittaker, 2016, amongst many). Further work should seek to analyse within-ball variation in greater depth. Incorporating individual level characteristics in a multilevel TDA framework that allows for regional context would also be valuable. Despite widespread use in the literature, the necessary reliance on 2011 Census data to understand the referendum outcome raises questions for additional research. Notwithstanding these opportunities for further work, the results presented in this paper represent a robust visualisation of the economic geography of Brexit, reflecting a multifaceted narrative of the build up to the referendum and subsequent policy response.

DISCLOSURE STATEMENT

FUNDING

and mutually funded by the Polish Ministry of Science and Higher Education and the German Federal Ministry of Education and Research.

NOTES

1. In this paper we define 'Leave' as a contraction of 'the UK should leave the EU'; and 'Remain' as the 'UK should remain in the EU'. In keeping with the literature 'Britain' and 'UK' are used interchangeably.

2. Issues of policy mandate and levelling up are discussed by McCann and Ortega-Argilés (2021). See also the analysis of the distribution of funds through the Town Deals scheme according to 2019 political support (Hanretty, 2021).

3. Northern Ireland has been shown to behave similarly to Scotland, but is excluded from our analysis because of the unique political situation there.

4. A brief exposition of the approach follows in section 4. A fuller discussion is then given in the supplemental data online.

5. On EU exposure through trade and on immigration, see, for example, Los et al. (2017), Clarke and Whittaker (2016) and Goodwin and Milazzo (2017). On austerity policies, see Feltzer (2019). For earlier discussions of demography and education, see Clarke and Whittaker (2016) and Sampson (2017); along with the role of cultural values in Arnorsson and Zoega (2018).

6. Post-industrial regions may have similar demographics, but can be found in the North East, North West, Wales, Scotland and the West Midlands. Equally, all those regions contain growing cities such as Manchester, Leeds, Cardiff, Glasgow and Birmingham where new industry flourishes.

7. That is, that spatially aggregated data do not allow neighbourhood effects to be distinguished from individual-level effects. For a discussion of identification issues and the neighbourhood scale, see, for example, Johnston et al. (2004) and Knies et al. (2021).

8. For this reason we exclude party affiliations as axes in our main TDABM plot, though including them in our sensitivity analysis. Our conclusions are robust, perhaps because – to the extent that it was informative on the Brexit vote – party affiliation tends itself to be a product of the socio-economic characteristics already accounted for in the analysis (see the Appendix in the supplemental data online).

9. Analysis of the TDABM plot coloured by political party affiliation is available from the authors upon request.

10. Shaw et al. (2017) also note that 'Both sides … overtly pitched their argument to hit the emotions of individuals – less often geared towards a positive emotion' (p. 1027).

11. Intuitively, consider the Anscombe's quartet (Anscombe, 1973) used in the teaching of ordinary least squares (OLS) regression to remind students of the importance of visualising data. Anscombe presents four examples of bivariate data in which the estimated OLS model linking the independent variable X to the dependent variable Y is identical. In the four cases the X coordinates of the points differ, including an example to illustrate high leverage points in which the X distribution is twin-peaked. In the context of the dataset analysed in this paper, the estimated Leave vote is a function of many of the independent variables used in the analysis. Anscombe's results remind us that there is no influence from the existence of a model explaining Leave to the appearance of the dependent variables across the joint distribution of the constituency characteristics. The relationship between constituency characteristics and the Leave vote in the present literature does not dictate the Leave homogeneity and Remain heterogeneity evidenced in this paper.

12. Data are accessed from https://www.pippanorris.com/data/.

13. Election data are downloaded from https://commonslibrary.parliament.uk/research-briefings/cbp-8647/\#fullreport/.

14. For full summary statistics, see the supplemental data online.

15. To understand the intuition for this, consider the tenure variable percentage living rent-free. This has a minimum value close to 0% and a maximum close to 4%. A radius of 2 could then include all values of rent-free living if the values on every other axis were the same. This then extends to cases where the radius is much larger, such as the 23 used in this paper, where the whole range of rent free can be covered and still allow variation in other axes.

16. For a full exposition of the methodology and discussion of evaluating outcomes, see the supplemental data online.

17. Including ball 8 the average rises to 13.8.

18. For TDABM plots coloured by region, see the supplemental data online.

19. The dominant group within this category is students in ball 31, and lone resident and lone parent in ball 32. However, the differences are small in absolute terms because of the low proportion of households in each set.

20. The minority of Remain-voting Welsh constituencies are found in 24, 15 and 33; other Welsh constituencies cluster in the main Leave-leaning balls, for example, 23 and 1.

21. For a table of summary statistics for each of the options within the Census 2011 questions, see the supplemental data online. Summaries for the other robustness tests are available from the authors upon request.

22. Balls registering a low percentage in Figure 1(a, b) could signify either 'Conservative throughout' or 'not Conservative throughout'. Further plots analysing support for other parties are available from the authors upon request.

23. For example, Walsall North, Mansfield, Great Grimsby and Stoke-on-Trent in ball 1, Wrexham in ball 2, and Burnley in ball 36.

24. For full details of the variables included in each dataset, see the supplemental data online.

25.   While the BM algorithm has assigned different numbers to the balls in Figure 4(a) relative to Figure 1, the constituencies within the large, highly connected balls making up the Leave section of the plot are consistent with the earlier plot. Lists of constituencies are available from the authors upon request.

26.   For details of the technical process used, see Hanretty and Vivyan (2015).

ORCID

Simon Rudkin http://orcid.org/0000-0001-8622-7318
Lucy Barros http://orcid.org/0000-0002-8421-107X
Paweł Dłotko http://orcid.org/0000-0001-5352-3102
Wanling Qiu http://orcid.org/0009-0006-1044-2568

REFERENCES

Abreu, M., & Öner, Ö. (2020). Disentangling the Brexit vote: The role of economic, social and cultural contexts in explaining the UK's EU referendum vote. *Environment and Planning A: Economy and Space*, *52*(7), 1434–1456. https://doi.org/10.1177/0308518X20910752

Alabrese, E., Becker, S. O., Fetzer, T., & Novy, D. (2019). Who voted for Brexit? Individual and regional data combined. *European Journal of Political Economy*, *56*, 132–150. https://doi.org/10.1016/j.ejpoleco.2018.08.002

Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician*, *27*(1), 17–21. https://doi.org/10.1080/00031305.1973.10478966

Antonucci, L., Horvath, L., Kutiyski, Y., & Krouwel, A. (2017). The malaise of the squeezed middle: Challenging the narrative of the 'left behind' Brexiter. *Competition & Change*, *21*(3), 211–229. https://doi.org/10.1177/1024529417704135

Arnorsson, A., & Zoega, G. (2018). On the causes of Brexit. *European Journal of Political Economy*, *55*, 301–323. https://doi.org/10.1016/j.ejpoleco.2018.02.001

Ashcroft, M. (2019). How Britain voted and why: my 2019 General Election post-vote poll. *Lord Ashcroft Polls*, 13 December. https://lordashc roftpolls.com/2019/12/how-britain-voted-and-why-my-2019-general-election-post-vote-poll

Axe-Browne, A., & Hansen, M. (2021). Still dividing the electorate? Brexit and voter evaluation of candidates. *Journal of Elections, Public Opinion and Parties*, *31*(2), 220–242. https://doi.org/10.1080/17457289.2019.1711097

Barca, F., McCann, P., & Rodriguez-Pose, A. (2012). The case for regional development intervention: Place-based versus place-neutral approaches. *Journal of Regional Science*, *52*(1), 134–152. https://doi.org/10.1111/j.1467-9787.2011.00756.x

Becker, S. O., Fetzer, T., & Novy, D. (2017). Who voted for Brexit? A comprehensive district-level analysis. *Economic Policy*, *32*(92), 601–650. https://doi.org/10.1093/epolic/eix012

Billing, C., McCann, P., & Ortega-Argilés, R. (2019). Interregional inequalities and UK sub-national governance responses to Brexit. *Regional Studies*, *53*(5), 741–760. https://doi.org/10.1080/00343404.2018.1554246

Billing, C., McCann, P., Ortega-Argilés, R., & Sevinc, D. (2021). UK analysts' and policy-makers' perspectives on Brexit: Challenges, priorities and opportunities for subnational areas. *Regional Studies*, *55*(9), 1571–1582. https://doi.org/10.1080/00343404.2020.1826039

Bromley-Davenport, H., MacLeavy, J., & Manley, D. (2018). Brexit in Sunderland: The production of difference and division in the UK referendum on European Union membership. *Environment*

and Planning C: Politics and Space. https://doi.org/10.1177/0263774X18804225

Clarke, H., Goodwin, M., & Whiteley, P. (2017). Why Britain voted for Brexit: An individual-level analysis of the 2016 Referendum vote. *Parliamentary Affairs*, *70*, 439–464. https://doi.org/10.1093/pa/gsx005

Clarke, S., & Whittaker, M. (2016). *The importance of place: Explaining the characteristics underpinning the Brexit vote across different parts of the UK*. Resolution Foundation.

Cooper, L., & Cooper, C. (2021). 'Get Brexit done': The New political divides of England and Wales at the 2019 election. *The Political Quarterly*, *91*(4), 751–761. https://doi.org/10.1111/1467-923X.12918

Cutts, D., Goodwin, M., Heath, O., & Surridge, P. (2020). Brexit, the 2019 general election and the realignment of British politics. *The Political Quarterly*, *91*(1), 7–23. https://doi.org/10.1111/1467-923X.12815

De Vries, C. E., Van der Brug, W., Van Egmond, M. H., & Van der Eijk, C. (2011). Individual and contextual variation in EU issue voting: The role of political information. *Electoral Studies*, *30*(1), 16–28. https://doi.org/10.1016/j.electstud.2010.09.022

Dijkstra, L., Poelman, H., & Rodríguez-Pose, A. (2020). The geography of EU discontent. *Regional Studies*, *54*(6), 737–753. https://doi.org/10.1080/00343404.2019.1654603

Dłotko, P. (2019a). Ball Mapper: a shape summary for topological data analysis. *arXiv preprint arXiv:1901.07410*.

Dłotko, P. (2019b). *BallMapper: Create a BallMapper graph of the input data*. R package version 0.1.0.

Dłotko, P., Qiu, W., & Rudkin, S. (2021). Financial ratios and stock returns reappraised through a topological data analysis lens. *The European Journal of Finance, Pages*, 1–25. https://doi.org/10.1080/1351847X.2021.2009892

Feltzer, T. (2019). Did austerity cause Brexit? *American Economic Review*, *109*(11), 3849–3886. https://doi.org/10.1257/aer.20181164

Goodhart, D. (2016). *The road to somewhere: The New tribes shaping British politics*. Penguin.

Goodwin, M., & Heath, O. (2016a). *Brexit vote explained: poverty, low skills, and lack of opportunities* (Report). Joseph Rowntree Foundation.

Goodwin, M., & Heath, O. (2016b). The 2016 referendum, Brexit and the left behind: An aggregate-level analysis of the result. *The Political Quarterly*, *87*(3), 323–332. https://doi.org/10.1111/1467-923X.12285

Goodwin, M., Hix, S., & Pickup, M. (2018). For and against Brexit: A survey experiment of the impact of campaign effects on public attitudes toward EU membership. *British Journal of Political Science*, *50*, 481–495. https://doi.org/10.1017/S0007123417000667

Goodwin, M., & Milazzo, C. (2017). Taking back control? Investigating the role of immigration in the 2016 vote for Brexit. *British Journal of Politics and International Relations*, *19*, 450–464. https://doi.org/10.1177/1369148117710799

Gordon, I. (2018). In what sense left behind by globalisation? Looking for a less reductionist geography of the populist surge in Europe. *Cambridge Journal of Regions, Economy and Society*, *11*, 95–113. https://doi.org/10.1093/cjres/rsx028

Hanretty, C. (2017). Areal interpolation and the UK's referendum on EU membership. *Journal of Elections, Public Opinion and Parties*, *27*(4), 466–483. https://doi.org/10.1080/17457289.2017.1287081

Hanretty, C. (2021). The pork barrel politics of the towns fund. *Political Quarterly*, *92*(1), 7–13. https://doi.org/10.1111/1467-923X.12970

Hanretty, C., & Vivyan, N. (2015). *Estimating constituency opinion in Britain* (Technical Report).

Harris, R., & Charlton, M. (2016). Voting out of the European union: Exploring the geography of leave. *Environment and Planning A: Economy and Space*, *48*(11), 2116–2128. https://doi.org/10.1177/0308518X16665844

Hayton, R. (2022). Brexit and party change: The conservatives and labour at Westminster. *International Political Science Review, Pages*, *43*(3). https://doi.org/10.1177/01925121211003787

Heath, O., & Goodwin, M. (2017). The 2017 general election, Brexit and the return to two-party politics: An aggregate-level analysis of the result. *The Political Quarterly*, *88*(3), 345–358. https://doi.org/10.1111/1467-923X.12405

Henderson, A., Poole, E. G., Jones, R. W., Wincott, D., Larner, J., & Jeffery, C. (2021). Analysing vote-choice in a multinational state: National identity and territorial differentiation in the 2016 Brexit vote. *Regional Studies*, *55*(9), 1502–1516. https://doi.org/10.1080/00343404.2020.1813883

Hobolt, S. B. (2007). Campaign information and voting behaviour in EU referendums. In C. H. Vreese (Ed.), *The dynamics of referendum campaigns* (pp. 84–114). Springer.

Hobolt, S. B. (2016). The Brexit vote: A divided nation, a divided continent. *Journal of European Public Policy*, *23*(9), 1259–1277. https://doi.org/10.1080/13501763.2016.1225785

House of Commons. (2022). *Local economic growth: Fifth Report of Session 2022–23*. Committee of Public Accounts.

Inglehart, R. F., & Norris, P. (2016). *Trump, Brexit, and the rise of populism: Economic have-nots and cultural backlash*.

Johnston, R., Jones, K., Sarker, R., Propper, C., Burgess, S., & Bolster, A. (2004). Party support and the neighbourhood effect: Spatial polarisation of the British electorate, 1991–2001. *Political Geography*, *23*, 367–402. https://doi.org/10.1016/j.polgeo.2003.12.008

Johnston, R., Manley, D., Pattie, C., & Jones, K. (2018). Geographies of Brexit and its aftermath: Voting in England at the 2016 Referendum and the 2017 general election. *Space and Polity*, *22*(2), 162–187. https://doi.org/10.1080/13562576.2018.1486349

Knies, G., Melo, P., & Zhang, M. (2021). Neighbourhood deprivation, life satisfaction and earnings: Comparative analyses of neighbourhood effects at bespoke scales. *Urban Studies*, *58*(13), 2640–2659. https://doi.org/10.1177/0042098020956930

Lee, N., Morris, K., & Kemeny, T. (2018). Immobility and the Brexit vote. *Cambridge Journal of Regions, Economy and Society*, *11*, 43–163. https://doi.org/10.1093/cjres/rsx027

Liberini, F., Oswald, A. J., Proto, E., & Redoano, M. (2019). Was Brexit triggered by the old and unhappy? Or by financial feelings? *Journal of Economic Behavior & Organization*, *161*, 287–302. https://doi.org/10.1016/j.jebo.2019.03.024

Los, B., McCann, P., Springford, J., & Thissen, M. (2017). The mismatch between local voting and the local economic consequences of Brexit. *Regional Studies*, *51*(5), 786–799. https://doi.org/10.1080/00343404.2017.1287350

McCann, P. (2020). Perceptions of regional inequality and the geography of discontent: Insights from the UK. *Regional Studies*, *54*(2), 256–267. https://doi.org/10.1080/00343404.2019.1619928

McCann, P., & Ortega-Argilés, R. (2021). The UK 'geography of discontent': Narratives, Brexit and inter-regional 'levelling up'. *Cambridge Journal of Regions, Economy and Society*, *14*, 545–564. https://doi.org/10.1093/cjres/rsab017

Mckenzie, L. (2016). *Brexit is the only way the working class can change anything*.

Qiu, W., Rudkin, S., & Dłotko, P. (2020). Refining understanding of corporate failure through a topological data analysis mapping of Altman's Z-score model. *Expert Systems with Applications*, *156*, 113475. https://doi.org/10.1016/j.eswa.2020.113475

Sampson, T. (2017). Brexit: The economics of international disintegration. *Journal of Economic Perspectives*, *31*(4), 163–184. https://doi.org/10.1257/jep.31.4.163

Shaw, D., Smith, C. M., & Scully, J. (2017). Why did Brexit happen? Using causal mapping to analyse secondary, longitudinal data. *European Journal of Operational Research*, *263*, 1019–1032. https://doi.org/10.1016/j.ejor.2017.05.051

The Guardian. (2019a). Guardian view on general election 2019: A fleeting chance to stop Boris Johnson in his tracks. *The Guardian*, 10 December, 19:13 GMT. https://www.theguardian.com/commentisfree/ng-interactive/2019/dec/10/the-guardian-view-on-general-election-2019-a-fleeting-chance-to-stop-boris-johnson-in-his-tracks

The Guardian. (2019b). Guardian view on the 2019 election result: a new political landscape. *The Guardian*, 13 December, 17:32 GMT. https://www.theguardian.com/commentisfree/2019/dec/13/the-guardian-view-on-the-2019-election-result-a-new-political-landscape

Thorsen, E., Jackson, D., & Lilleker, D. (2017). *UK election analysis 2017: Media, voters and the campaign*. Centre for the Study of Journalism, Culture & Community.

van der Merwe, B., & Goodier, M. (2022). *How Tory seats hoovered up levelling-up funds*. 15 June, updated 5 August.

Wincott, D. (2021). Multilevel dynamics of the UK's EU exit: Identity, territory, policy and power. *Regional Studies*, *55*(9), 1491–1501. https://doi.org/10.1080/00343404.2021.1950915

Zhang, A. (2018). New findings on key factors influencing the UK's referendum on leaving the EU. *World Development*, *102*, 304–314. https://doi.org/10.1016/j.worlddev.2017.07.017