Inference and Influence of Network Structure Using Snapshot Social Behaviour without Network Data

(Supplementary Materials)

Antonia Godoy-Lorite, Nick S. Jones

1 Datasets

The electoral outcomes we use in Section IV in the main text are London Mayoral Election 2012 and 2016 from http://data.london.gov.uk/elections, and the EU Referendum for Greater London from the BBC's electoral wards outcomes at https://3859gp38qzh51h504x6gvv0o-wpengine.netdna-ssl.com/files/2017/02/ward-results.xlsx. For Mayoral Elections we only use Conservative and Labour votes to keep to binary outcomes, since they account for the 84% of 2012 ME and 79% 2016 ME. We use data at the electoral ward level, given that it is the smallest spatial resolution that we get for electoral outcomes, such that Greater London is divided into 630 electoral wards with an average electorate of 7750 ± 1550 each. The EU Referendum ward level data are incomplete, there are 18 boroughs (out of 32 boroughs excluding the City of London) that did not release the outcomes at the electoral wards level but aggregated them at Borough level, which are: Bexley, Brent, Barnet, Barking and Dagenham, Newham, Tower and Hamlets, Sutton, Hillingdon, Lewisham, Redbridge, Southwark, Wandsworth, Hackney, Kingston, Richmond, Hammersmith and Fulham, Kensington and Chelsea, and Westminster. For those Boroughs we are using the aggregated outcomes at Borough level at http://data.london.gov.uk/electi ons, also for 2012 and the 2016 London Mayoral Elections in Figs.S9-S6. Regarding the Blau space coordinates of each ward, we compute the average for each of the chosen dimension of the Blau space, which are: education (categorical variable with 5 categories), age (ordinal variable in binning of 1 year from 18 to 100 years), gender (binary variable 0 for male and 1 for female) all of them from census data 2011 https://www.ons.gov.uk/census/2011census, http://infusecp.mimas.ac.uk/, then spatial location (centroid coordinates of each ward) and income in 1000GBP (median household income estimates from 2012/2013 https://data.london.gov .uk/dataset/ward-profiles-and-atlas). Selected variables were chosen based on availability from official sources, comparability to the homophilic features in McPherson and Smith (2019) and Hoffmann and Jones (2020), and from computational limitations when including variables with multiple categories that require an EF parameter for each (as for example race or religion). The Blau space coordinates are not standardised –only the relative distances are standardised– but we use age and income coordinates multiplied by a factor 0.1 and gender by a factor 10 so that they are of the same order of magnitude. The City of London has a special treatment inside Greater London, it is not considered a Borough as the other 32 Boroughs in GL and there is no census data on it, thus we did not include it in our analysis. Also due to a miss-match between the definition of electoral wards used for electoral outcomes and 2011 census wards, we could only use 608 wards out of the 630 electoral wards and the 625 2011 census wards. All raw and clean data are publicly available in https://github.com/agodoylo/kernel-Blau-Isingmodel (DOI: 10.5281/zenodo.4336456).

2 Extensions of the KBI approach

2.1 Potts model extension

In this section we discuss the extension from a $q = 2$ Ising model to $q > 2$ Potts model of out KBI approach. First, the *generative model for social structure* does apply to the Potts model: we have N individuals embedded in a K-dimensional Blau space with coordinates $\boldsymbol{z_i} \in R^K$; and individuals are connected according to their distances in the Blau space with probability from Eq. (2) in the main-text.

As for the *behavioural model*, in the Potts model each individual spin σ_i can take more than 2 directions $\sigma_i = \{1, .., q\}$ with $q > 2$. As in the original KBI, the spin orientation depends on the external fields and the spins they are connected in the network. Using linear fields in the Blau space, while for $q = 2$ (Ising model) the interaction with the external fields is the scalar product $\boldsymbol{h} \cdot \boldsymbol{z_i} = \sum_k h_k z_{ik}$, for $q > 2$ there are $q' = \{1, .., q\}$ sets of linear fields $h_k^{q'}$, each one governing each of the q possible spins direction. The interaction with the external fields for the Potts extension in each q' direction is,

$$(\boldsymbol{h} \cdot \boldsymbol{z_i})_{q'} = \sum_k h_k^{q'} z_{ik}.$$

Finally, the energy of a spin configuration σ is given by the Hamiltonian function,

$$H(\boldsymbol{\sigma}, \boldsymbol{h}, J, \boldsymbol{A}) = -\sum_i \left(\sum_{q'} (\sum_k h_k^{q'} z_{ik}) \delta(\sigma_i, \sigma_{q'}) \right) - J \sum_{ij} A_{ij} \delta(\sigma_i, \sigma_j),$$

where $\sigma_{q'}$ is a 'ghost spin' that controls the contribution to the q' direction. Then, the configuration probability is given by the Boltzmann equation as for the KBI model and the same ABC inference method can be applied. However, although the kernel parameters remain the same, the number of parameters increases because of the external fields, and presumably the generation of spins configuration for $q > 2$ given a set of parameters $\Theta = \{\beta, \boldsymbol{h}, J, \boldsymbol{\theta}\}$ is going to be computationally more expensive.

2.2 Multiple observations extension

In this section we discuss the extension of the KBI model in the case where multiple observations of the same behaviours on the same population are available. In this case, we assume the observations are independent samples of the equilibrium, otherwise the model will require more modifications at the parametrisation level to capture, i.e. temporal trends. We have m

independent observations of the same behaviour on the same population $\{\boldsymbol{\sigma}^{\mathbf{1}}, ..., \boldsymbol{\sigma}^{\boldsymbol{m}}\}$, where $\boldsymbol{\sigma}^i \in \{-1, 1\}^N$ is the population spin configuration of the i^{th} observable. We can apply the same ABC inference framework as in Algorithm 1 in the main-text by extending the overall distance to the m observed configurations in Eq. 6 in the main-text as,

$$\tilde{\eta} = \frac{1}{m} \sum_{m} \eta(\boldsymbol{S}(\boldsymbol{\sigma}'^{\boldsymbol{m}}), \boldsymbol{S}(\boldsymbol{\sigma}^{\boldsymbol{m}})) = \frac{1}{m} \frac{1}{N} \sum_{m} \sum_{z=1}^{C} n_z |S_z'^m - S_z^m|. \tag{S1}$$

3 Inferring model parameters in Synthetic data

3.1 Rescaling population size

For computational expedience our treatment of electoral data uses a subsample of the full population. Here we make the case for approximating systems of size $N^{(1)}$ with smaller systems of size $N^{(2)}$. We suggest that inference of the kernel parameters (except the bias parameter θ_0) can be insensitive to changes in the population size N. We demand that the update dynamics of spin i be conserved under system-size rescaling where, each Blau co-ordinate z, has spin-up fraction S_z. The external field's effect on each spin's update dynamics does not change when the system size is rescaled and so can be neglected. Single-spin update dynamics rescaling thus depends only on the term $\sum_j A_{ij}\sigma_j$: it suffices to show that this term is unchanged under a system size rescaling (when combined with a rescaling of θ_0).

If the number of spins at each Blau-space co-ordinate is sufficiently large then $\sum_j A_{ij}\sigma_j \sim \sum_z \rho_{iz} S_z n_z$ where n_z is the number of spins at co-ordinate z and ρ_{iz} is the chance of connecting i to any node at location z. If we require that when we approximate systems of size $N^{(1)}$ with smaller systems of size $N^{(2)}$ we have: a) $n_z^{(1)}$ and $n_z^{(2)}$ are related by $\frac{n_z^{(1)}}{n_z^{(2)}} = \frac{N^{(1)}}{N^{(2)}}$ (i.e. we rescale the population size of each Blau-space co-ordinate in a manner proportional to the overall system-size) and b) that $\theta_0^{(2)} = \theta_0^{(1)} - log(N^{(1)}/N^{(2)})$, $\theta_k^{(2)} = \theta_k^{(1)}$ $k \neq 0$, then we find that $\sum_z \rho_{iz}^{(1)} S_z n_z^{(1)} = \sum_z \rho_{iz}^{(2)} S_z n_z^{(2)}$. This result follows straightforwardly given requirement a) and because requirement b) ensures $\rho_{iz}^{(2)} = \frac{N^{(1)}}{N^{(2)}} \rho_{iz}^{(1)}$. Given that $\sum_z \rho_{iz}^{(1)} S_z n_z^{(1)} = \sum_z \rho_{iz}^{(2)} S_z n_z^{(2)}$ the update dynamics of each spin are unchanged under rescaling.

We now justify why $\theta_0^{(2)} = \theta_0^{(1)} - log(N^{(1)}/N^{(2)})$ implies $\rho_{iz}^{(2)} = \frac{N^{(1)}}{N^{(2)}} \rho_{iz}^{(1)}$. A_{ij} is Bernoulli distributed according to the kernel $\rho(i,j) = \frac{1}{1+exp(\theta_0 + \sum_k \theta_k |z_{ik} - z_{jk}|)}$. For $N^{(1)} \gg 1$ and $N^{(2)} \gg 1$, but small finite expected degree, the connection probabilities are low so that $\rho(i,j) = \frac{1}{1+exp(\theta_0 + \sum_k \theta_k |z_{ik} - z_{jk}|)} \approx e^{-\theta_0} e^{-\sum_k \theta_k |z_{ik} - z_{jk}|}$, therefore if $\theta_0^{(2)} = \theta_0^{(1)} - log(N^{(1)}/N^{(2)})$, then $\rho^{(2)} \approx \frac{N^{(1)}}{N^{(2)}} e^{-\theta_0^{(1)}} e^{-\sum_k \theta_k^{(1)} |z_{ik} - z_{jk}|} \approx \frac{N^{(1)}}{N^{(2)}} \rho^{(1)}$.

As seen in Fig. S1 the approximation yields similar inference results, as regards the maxima (and general form) of the posteriors. A consequence of using a smaller system to represent a larger is that the smaller system can show proportionately larger fluctuations; this can inflate the variance of the posteriors leading to an overestimate in parametric uncertainty. A comparison of the right and left sides of Fig. S1 shows that the smaller system has slightly wider posteriors. While suitable for our illustrative objective here, this is a direction that requires further optimisation should sharper posteriors be required.

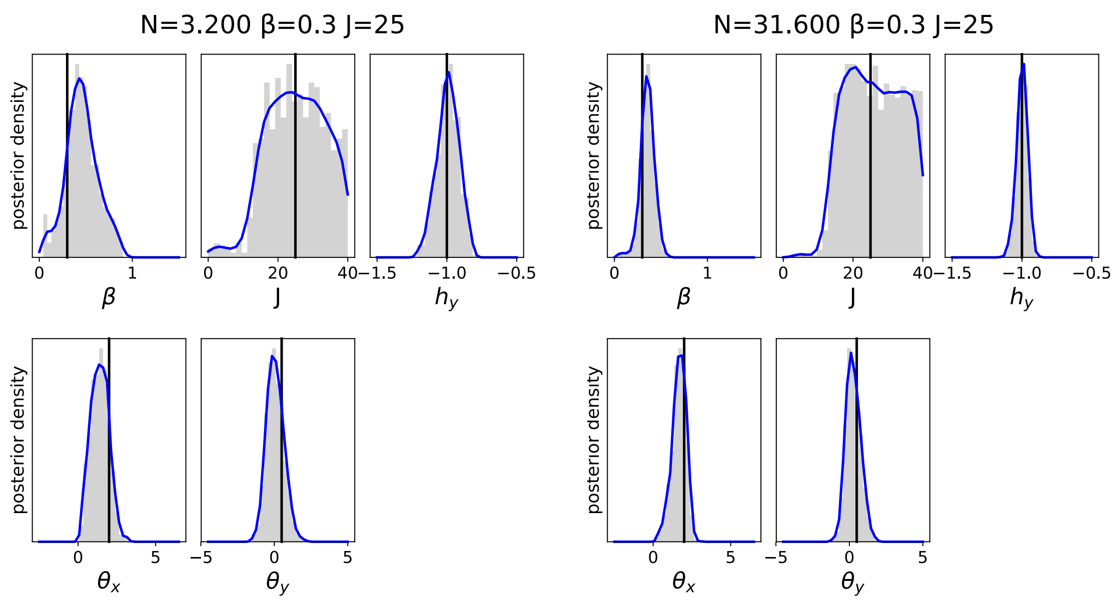

Figure S1: **Dimension-specific kernel parameters are insensitive to a rescaling of the system size**. As input data for the inference we use a single-observation of the summary outcomes S_z for $z = (x, y) \in x, y = \{0, 1, .., 9\}$ generated for $N = 10,000$ for Fig. 2 in the main text with $\theta_0 = 9$ (only for $J = 25$). We show the ABC marginal posteriors for ≈ 500 samples from a sampling procedure of over $6M$ i.i.d samples for two different system sizes: on the left figure we show the ABC marginal posteriors for a smaller system with $N = 3,200$ ($\theta_0 = 7.86$, $= 0.055$) and on the right for a larger system with $N = 31,600$ spins ($\theta_0 = 10.15$, $\epsilon = 0.025$). We show that given the same single-observable outcomes we can re-scale the system size and recover indicative model parameters, and in particular the dimension-specific kernel parameters which are insensitive to the re-scaling.

3.2 Degeneracy between network degree and the connection strength

To illustrate the degeneracy between the bias term θ_0 and the connection strength in the Hamiltonian (Eq.3 in the main text), we perform the inference of the model parameters for an observable generated with $\theta_0 = 9$ ($\langle\kappa\rangle \approx 2$) but now sampling spin configurations with $\theta_0 = 7$ ($\langle\kappa\rangle \approx 10$) for a much higher density of connections.

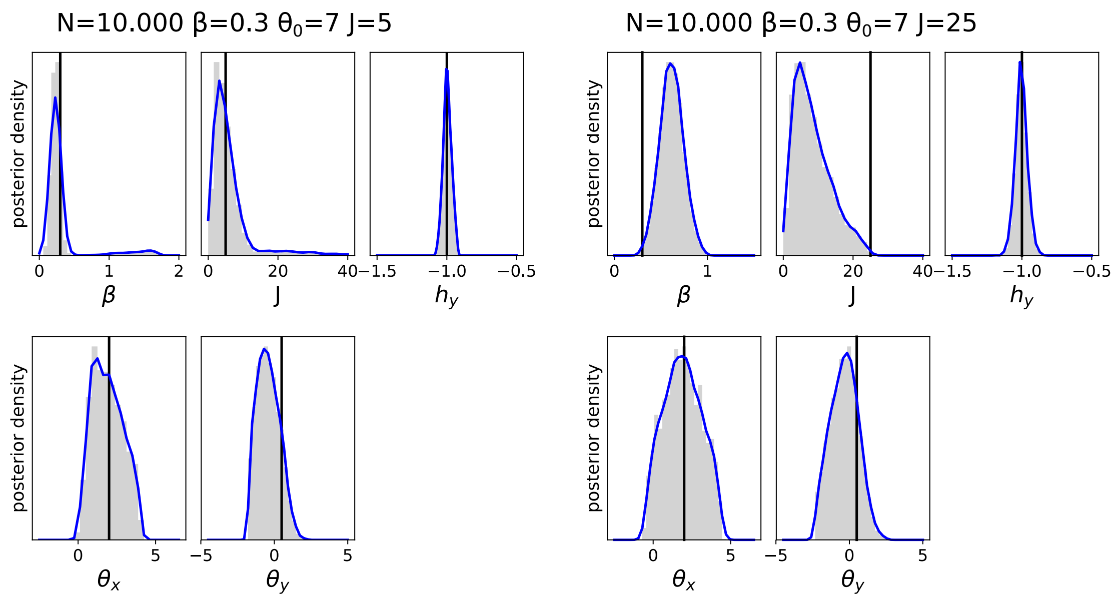

Figure S2: **Dimension-specific kernel parameters are insensitive to changes in the degree of the network**. As in Fig. S1, we use as input data for the inference the single-observation of the summary outcomes S_z for $z = (x, y) \in x, y = \{0, 1, .., 9\}$ generated for Fig. 2 in the main text (with $N = 10,000$ and $\theta_0 = 9$) for weak connection strength ($J = 5$) and strong connection strength ($J = 25$). We show the ABC marginal posteriors for ≈ 500 samples from a sampling procedure of over $6M$ i.i.d samples (under tolerance thresholds of $\epsilon = 0.045$ for $J = 5$ and $\epsilon = 0.037$ for $J = 25$), assuming a higher connectivity density $\theta_0 = 7$, corresponding to an average degree of $\langle\kappa\rangle \approx 10$. We found that the connection strength decreases when increasing the density of connections. However, although ABC marginal posteriors result in wider distributions, we are able to extract properties of the dimension-specific kernel parameters, θ_x, θ_y, from systems with higher connectivity density.

4 Synthetic and electoral data magnetisation

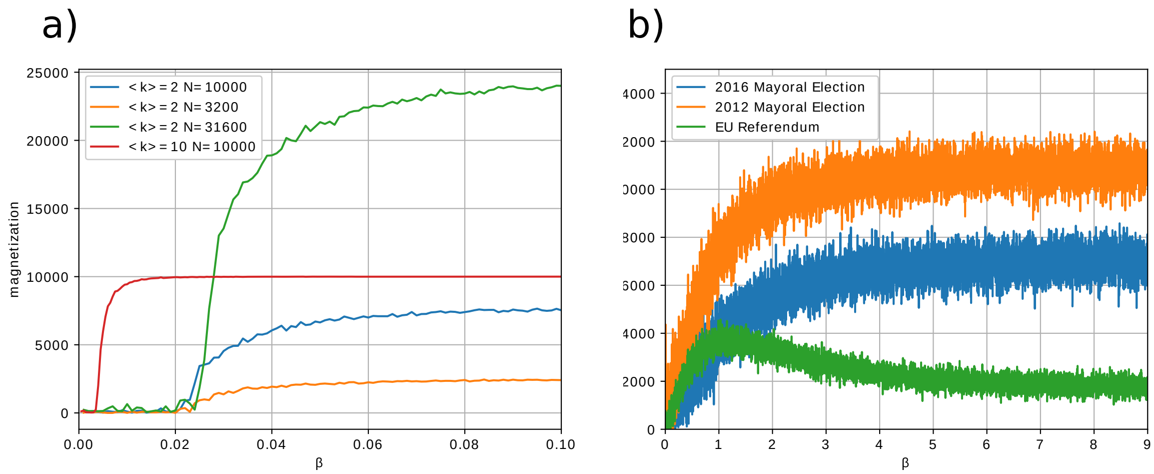

Figure S3: **Magnetisation both for synthetic data and for Greater London electoral data**. **a** We show the spins magnetisation on synthetic data for different system sizes and different connectivity density kernels (see the legend), but same EFs ($h_x = 1, :\ h_y = -1$). **b** Spins magnetisation for 2012 and 2016 London Mayoral Elections and EU Referendum. Except for β, the model parameters used to generate the different spin configurations are from a multivariate Gaussian fit on the ABC marginal posteriors in Fig. 4 in the main text.

5 Bootstrapping population Blau space coordinates from Census micro-data

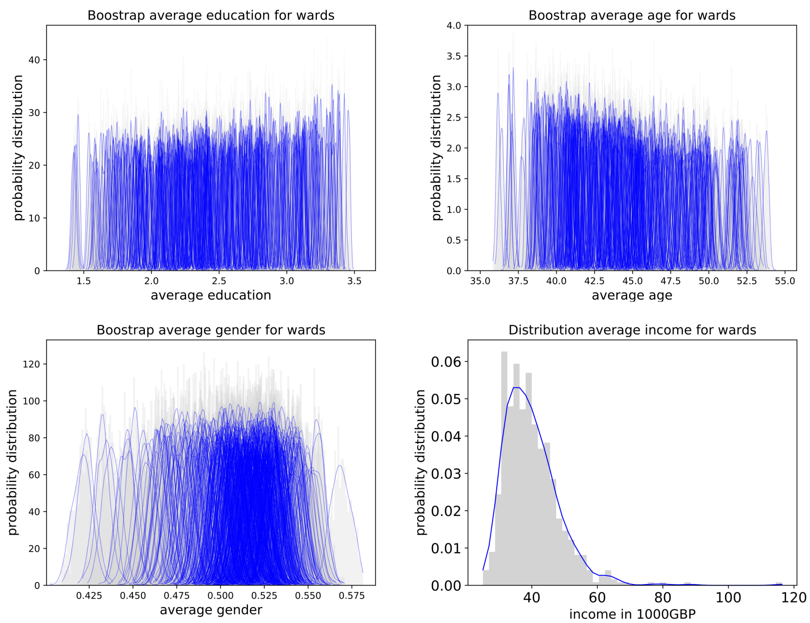

Figure S4: **Evidence of heterogeneous distribution of average wards' coordinates by bootstrapping census micro-data at the ward level**. The different sub-figures show the bootstrapping averages for the different 608 wards in Greater London from 2011 census data distribution on education, age and gender for $1,000$ samples of the actual ward population size (~ 7750 individuals). 2011 census data does not include an income variable, hence, we do not have ward level micro-data but only ward level income estimates (see Section VI in the main text), therefore, we show the overall probability distribution of income for the 608 wards in GL.

6 Multiple Linear Regression on electoral data

In order to get some simple estimates of what could be expected from a simple linear regression on the Blau space, we perform a multiple linear regression for 2012 and 2016 MEs and the EU Referendum electoral outcomes with education, age, gender and income as independent variables (we exclude geospatial (x, y) Blau coordinates from the linear regression). The Blau space coordinates are not standardised – only the distances are standardised– but we use age and income coordinates multiplied by a factor 10 and gender by a factor 0.1 so that they are of the same order of magnitude:

	education	age	gender	income
2012 ME	0.460	-0.333	0.489	-0.553
2016 ME	0.459	-0.491	0.583	-0.457
EU Referendum	0.453	-0.376	0.177	-0.033

7 ABC inference procedure for electoral data

In this section we detail the procedure we follow to obtain the ABC marginal posteriors in Fig. 4 in the main-text. First, on our choice of priors we used uniform priors for all model parameters since we did not have specific information on them but they were are all of similar order of magnitude (hence not using log-uniform priors). We needed to specify upper and lower limits to each prior, and, as we now detail, we set those limits by using prior literature values and a conservative previous exploration with the same ABC algorithm and larger ranges. We used pre-existing values and permissive fits of the model to establish a very approximate value of the parameters: multiple linear regression for the external fields (coefficients from the previous section SM ST6) and for the connectivity Kernel parameters we use as reference those from Hoffmann's homophilious kernels (Hoffmann and Jones 2020). Estimates of connectivity kernels parameters from Hoffmann and Jones (2020) are inferred from the British Household Panel Survey (BHPS), ego-network dataset collected at national level where respondents were asked about characteristics on "their closest friends", and they present values between 0 and 5. In this paper we use the same kernel function and same Blau space distance definition.

For the initial exploration of the parameter priors we set $\theta_k \in [-12, 14]$ for the kernel parameters and for the external fields parameters we choose as initial priors $h_{negative} \in [-4, 2]$ for the negative EFs from the multilinear regression, age and incone, and $h_{postive} \in [-2, 4]$ for the positive EFs in the linear regression, gender (education is set to 0.45). Then using the range of accepted values with a very permissive epsilon we yielded more specific ranges. There was only one scenario where we found that our posterior had significant mass at the boundary of the prior (this can be an indicator of having priors that are too narrow) and this was for the connectivity strength parameter J. However, in this case there were clear physical reasons why we could expect to see posteriors with mass at the boundary as we show in the main-text. This two step approach allowed us to return well-sampled posteriors by quickly ruling out large areas of parameter space: it can be interpreted as a simple one-step SMC ABC that allowed us to preserve the algorithmic clarity offered by rejection-based sampling while being computationally feasible.

Based on this procedure, for the final sampling in Fig. 4 in the main-text we set as priors: $\beta \in [0, 4]$, $J \in [0, 40]$, $h_{age} \in [-3, 0.15]$, $h_{gender} \in [-0.3, 3]$, $h_{income} \in [-1.6, 0.7]$, $\theta_{education} \in [-7, 12]$, $\theta_{age} \in [-5, 12]$, $\theta_{gender} \in [-6, 12]$, $\theta_{distance} \in [-7, 11]$ and $\theta_{income} \in [-5, 12]$. We sampled over $150M$ i.i.d samples for the three elections and save them. Finally, for the marginal ABC posteriors, we chose the tolerance threshold for each election a posteriori once all samples were generated. The tolerance thresholds are set by reducing ϵ until it does not have an effect on the posteriors, as is shown below in Section 8, while being close the the minimum distance generated by a very large sampling $\eta(S, S')_{min}$. From a theoretical range of distances between 0 and 1, we set the tolerance thresholds as,

	ϵ	$\eta(S, S')_{min}$
2012 Mayoral Election	0.088	0.081
2016 Mayoral Election	0.086	0.076
EU Referendum	0.050	0.044

8 ABC inference on electoral data for different tolerances ϵ

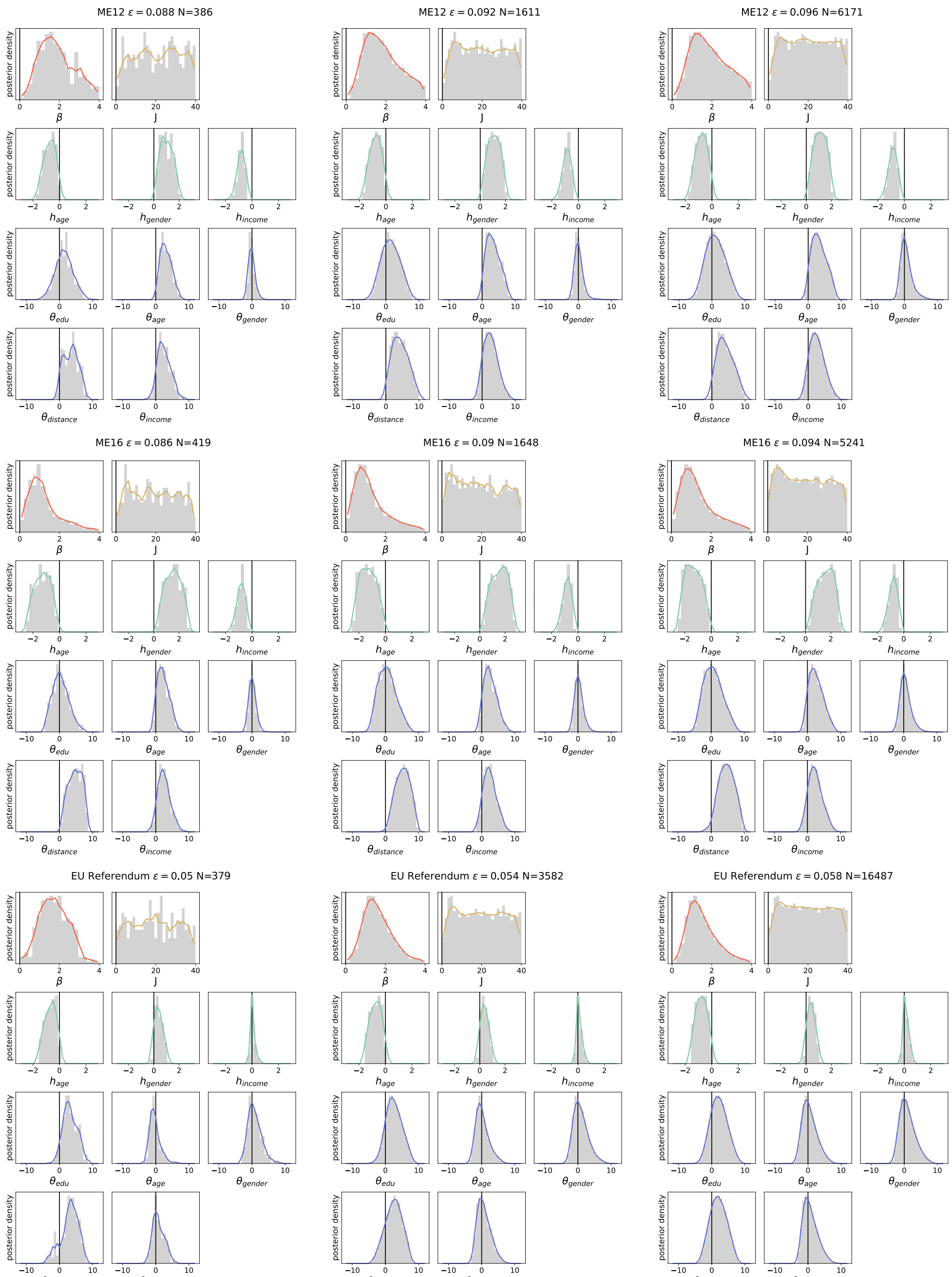

Figure S5: **ABC marginal posteriors for different tolerance thresholds** ϵ. As a complementary figure to Fig. 4 in the main text, here we show for each election the ABC parameters estimates for three different threshold tolerance. We show that from the largest tolerance, making the threshold tolerance smaller has no effect on the ABC marginal posteriors other than having slightly tighter posterior distributions.

9 ABC inference on electoral data with 280 wards and 18 boroughs as input data

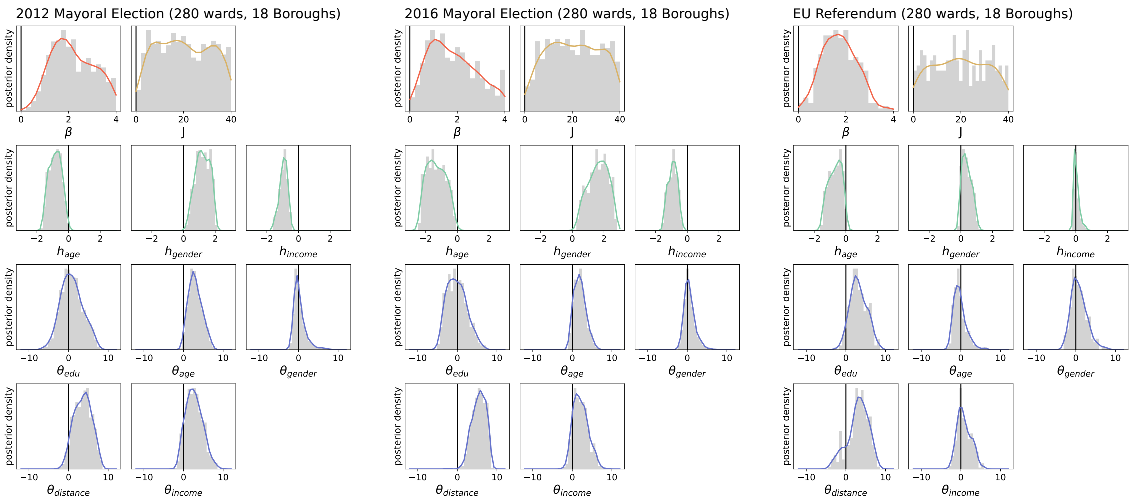

Figure S6: **ABC marginal posteriors for 2012 and 2016 London Mayoral Elections and the EU referendum using 280 electoral wards and 18 boroughs as input data**. As a complementary figure to Fig. 3 in the main text, here we show the ABC parameters estimates for the three elections with the same input data as EU Referendum data (280 ward outcomes and 18 borough outcomes). We show the ABC marginal posteriors for ≈ 500 samples (from over $150M$ i.i.d samples) under tolerance thresholds of $\epsilon = 0.071$ for 2012 ME, $\epsilon = 0.068$ for 2016 ME and and $\epsilon = 0.050$ for the EU Referendum. In the inference procedure, we choose to fix $\theta_0 = 14$ and $h_{edu} = 0.45$ given the degeneracies of the model (Eq.4 in the main text). We show the histograms of the ABC marginal posteriors in grey, as a solid line the Gaussian kernel density estimates, and the vertical lines correspond to 0 value. The parameters estimates are almost the same for these input data as for using 608 wards input data in Fig. 3 in the main text, with distance being the strongest homophilous signal and age and income also showing significant homiphilic signal.

10 Comparison of homophily kernel parameters inferred from the KBI model on Greater London elections with those inferred from ego-network survey data

The virtue of our inference method is that it attempts to infer social connectivity kernels that are specifically relevant for each particular behaviour (in the empirical examples in the main-text the behaviour is voting for different elections). In this section we compare the inferred homophily kernel parameters from the main-text with the closest available homophily-based network model parameters, which are homophilic kernels from ego-network survey data in "Inference of a universal social scale and segregation measures using social connectivity kernels" Hoffmann and Jones (2020). Hoffmann and Jones use data from the British Household Panel Survey (BHPS) (from 1992 to 2006 every other year) and from the Understanding Society Survey (USS) (2011 and 2014). In these surveys respondents are asked about their own socio-demographic characteristics and the socio-demographic characteristics of their 'closest friends'. This data is an appropriate comparison since the kernels are inferred from a real ego-network with information on both respondents and their closest friends. We note that a perfect comparator survey would have asked respondents not about their close friends, but instead, about the characteristics of people with whom they discuss topics regarding Mayoral Elections and the EU Referendum. We further note that the ego-networks survey data is for the whole UK and our experimental data are for Greater London only—kernel parameters may also differ between regions.

The homophilic kernels used in Hoffman and Jones (2020) and the standardisation of the Blau space distances are the same as the ones we used in our KBI model, which makes the parameters comparable. The inference procedure for both experiments are different, in accord with the different nature of the data. For the ego-networks the inference is based on a substantial amount of survey data (also including a bias term controlling edge density), thus the parameter errors are smaller than for our KBI parameters, where the inference does not include network data but we infer the network structure indirectly from observed behaviour (electoral outcomes). The next table shows, for each variable, the mean and standard deviations over all years in which ego-network surveys where collected (Hoffmann and Jones, 2020), and our inferred mean and standard deviation from the ABC marginal posteriors for the three elections in Fig. 4 in the main-text. Notice that education was not included in either of the UK surveys but in two ego-network surveys in the US: The American Life Panel (ALP) and The General Social Survey (GSS), which are also included in the table. The rest of the variables shown are from the two national UK ego-network surveys BHPS and USS; compared to our data they include occupation as a variable but do not include income. All data are available from the UK data service (http://doi.org/10.5255/UKDA-SN-6614-13), RAND American Life Panel (https://alpdata.rand.org/index.php?page=datap=showsurveysyid=86) and the US General Social Survey (https://gss.norc.org/). Code to reproduce results is available at https://github.com/tillahoffmann/kernels. Table S1 shows that uncertainty of the KBI model parameters is larger compared to ego-network-derived kernel parameters (they use different input data). However, the values are broadly consonant considering that the social context for the survey data (closest friends) is not the same as for our elections. Distance is the most relevant dimension segregating population for both experiments. Age is the next most relevant homophilic signal for ego-networks and Mayoral elections, but not for the EU referendum. We do not observe a significant homophily signal for gender in our KBI experiments, unlike the ego-network kernels, however, we believe this is due to the nature of our data: ward-level populations are well-mixed in gender, eliminating any possible signal. In general, we observe that the ego-networks based on closest friends are more aligned with Mayoral election kernels than with EU referendum kernel parameters, for which we observe education is a more segregating factor than age.

Table S1: Comparison of homophily kernel parameters inferred from our KBI model on Greater London elections with those inferred from ego-network survey data in the UK.

	ego-network	2012 ME	2016 ME	EU Referendum
Distance	2.83 ± 0.04	3.47 ± 2.16	4.75 ± 1.99	3.04 ± 2.53
Age	1.51 ± 0.08	3.10 ± 1.69	2.31 ± 1.71	-0.10 ± 1.59
Gender	1.34 ± 0.03	-0.09 ± 0.89	0.11 ± 0.93	0.87 ± 2.06
Education (ALP, GSS)	0.29 ± 0.06	0.98 ± 2.50	0.35 ± 2.43	3.12 ± 2.17
Occupation	0.47 ± 0.05	-	-	-
Income	-	2.57 ± 1.92	2.51 ± 1.83	0.84 ± 1.71

11 Maps of inferred electoral outcomes and connectivity networks

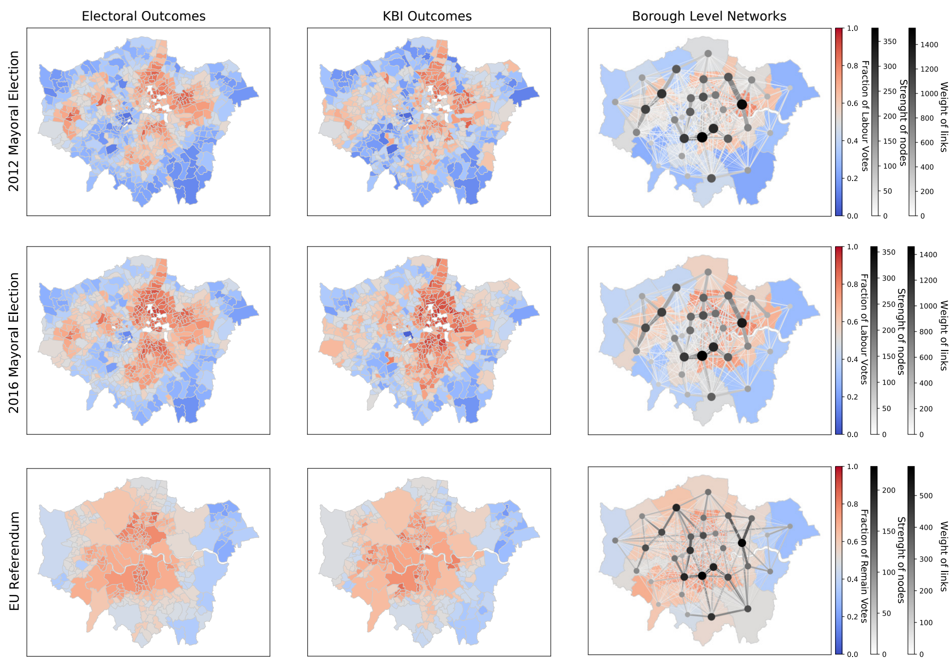

Figure S7: **KBI accurately reproduces electoral outcomes while inferred connectivity networks highlight differences between Mayoral Elections and EU referendum**. The *left* column show the real electoral outcomes and next to it the *middle* column show inferred outcomes generated from the best parameter set in the sampling procedure in Fig.4 in the main-text for the three election in Greater London. For London Mayoral elections, the colour range goes from red for $S_w = 1$ ward's fraction of Labour vote, and blue for $S_w = 0$ or all ward's votes for Conservatives. For EU Referendum, in order to keep colour consistency, we represent in red $S_{w/B} = 1$ fraction of ward/borough votes for remain, and blue $S_{w/B} = 0$ for wards/boroughs voting for leave. On the *right* column, we show the connectivity networks from Fig. 4 in the main-text ABC marginal posteriors of the kernel parameters aggregated at the borough level for visualisation purposes (ward level connections with $184,528$ links are hard to visualise). The colour map in the background is the real electoral outcomes at the borough level. For the connectivity kernel we coloured in a linear range (white to black) the strength of the nodes and the weight of the links, also the size of nodes and links is linear with their strength and weight respectively. We have not visualised intra-borough links but we have take them into account for the nodes strengths. We see that MEs connectivity networks are similar among them but different from the EU Referendum connectivity network. All the networks show pronounced spatial homophily leading to strong links between neighbouring boroughs.

12 ABC inference only with external fields ($J = 0$)

We show the ABC marginal posteriors for only external fields (without connectivity kernel) for different input data: for 2012 and 2016 London Mayoral Elections with 608 electoral wards outcomes; and given the missing wards' data for EU Referendum (see Section VI in main text), we use the available 280 ward level data and we use borough level electoral outcomes for the remaining 18 boroughs, for EU Referendum, and in order to compare the estimated model parameters with exactly the same input data, also for the 2012 and 2016 MEs.

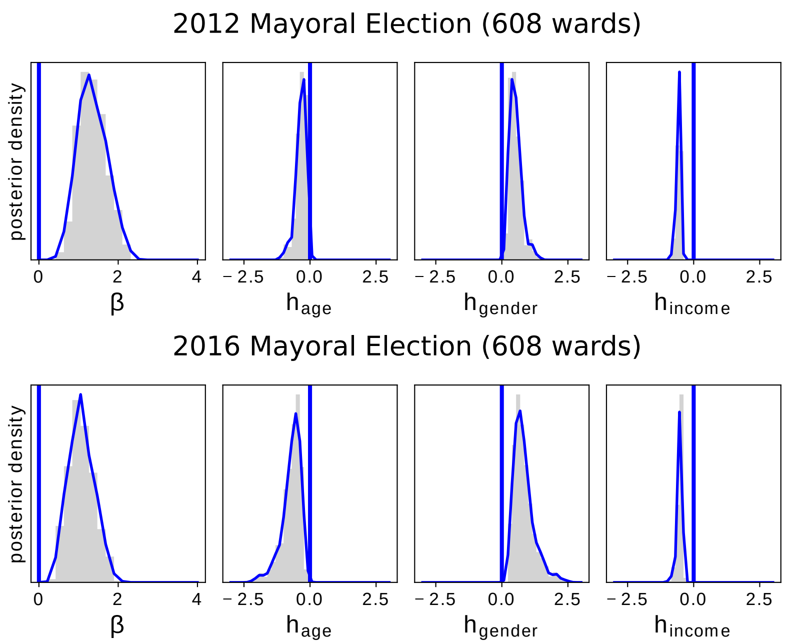

Figure S8: **Only external fields inference for 2012 and 2016 London Mayoral Elections with 608 electoral wards**. We show the ABC marginal posterior distribution for the model parameters excluding the connection term (only using linear external fields) for ≈ 500 samples from a sampling procedure of over $5M$ of i.i.d samples under tolerance thresholds of $\epsilon = 0.098$ for 2012 ME and $\epsilon = 0.095$ for 2016 ME. Given the degree of freedom in the Hamiltonian (Eq.3 in the main text), we choose to fix the education coefficient to $h_{edu} = 0.45$. We find that although 2012 ME external fields are peaked slightly closer to zero than 2016 ME external fields estimates, overall the estimates are similar for the two Mayoral Elections.

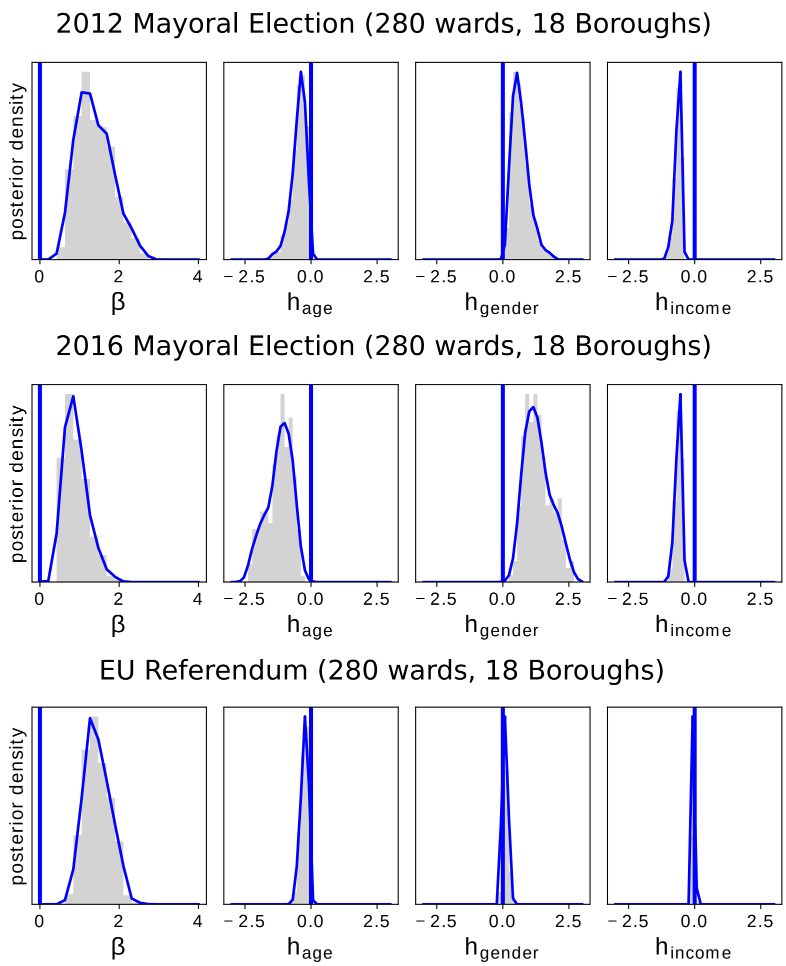

Figure S9: **Only external fields inference for 2012 and 2016 London Mayorial Elections and EU Referendum inference with 280 electoral wards and 18 boroughs**. We show the ABC marginal posterior distribution for the model parameters excluding the connection term (only using linear external fields) for ≈ 500 samples from a sampling procedure of over $5M$ of i.i.d samples under tolerance thresholds of $\epsilon = 0.083$ for 2012 ME, $\epsilon = 0.079$ for 2016 ME and $\epsilon = 0.052$ for the EU Referendum. Given the degree of freedom in the Hamiltonian (Eq.3 in the main text), we choose to fix the education coefficient to $h_{edu} = 0.45$. We find EFs estimates are more peaked around zero for the EU Referendum.

13 Electoral outcomes distributions and polarisation

In this section we show the polarisation of the different electoral ward outcomes distributions for the KBI generated outcomes with and without connections (connectivity kernel). We measure polarisation of wards outcomes as in J.M Esteban and D. Ray (2007)– without considering group identification– such as,

$$\mathscr{P}(S_w) = \frac{1}{N_L} \sum_{(i,j)\in L} |S_i - S_j|,$$

where S_i is the electoral outcome of ward i given as the fraction of votes to one of the two options, L and N_L are all pairs of electoral wards (for $i \neq$j) and the number of pairs of wards respectively. Notice the difference between measuring polarisation between wards' outcomes and polarisation inside wards or of the overall vote. Wards' outcomes polarisation would be zero if all wards have the same outcomes regardless if that is $S_w = 0,\ S_w = 1$ or $S_w = 0.5$, while overall votes polarisation would be 0 for $S_w = 0,\ S_w = 1$ and maximum for $S_w = 0.5$. We are focusing on the differences between ward-level outcomes as a proxy for the degree to which wards have similar voting composition to others, but not how far from consensus is the population vote. In what follows, we compare polarisation between the different electoral ward outcomes distributions from the ABC estimates in Fig. 3 in the main text, and for the generated outcomes with the same parameters but eliminating the connectivity kernel (with J=0).

Table S2: Electoral outcomes Polarisation

2012 ME	0.239	2012 ME with J=0	0.384
2016 ME	0.228	2016 ME with J=0	0.351
EU Referendum	0.163	EU Referendum with J=0	0.186

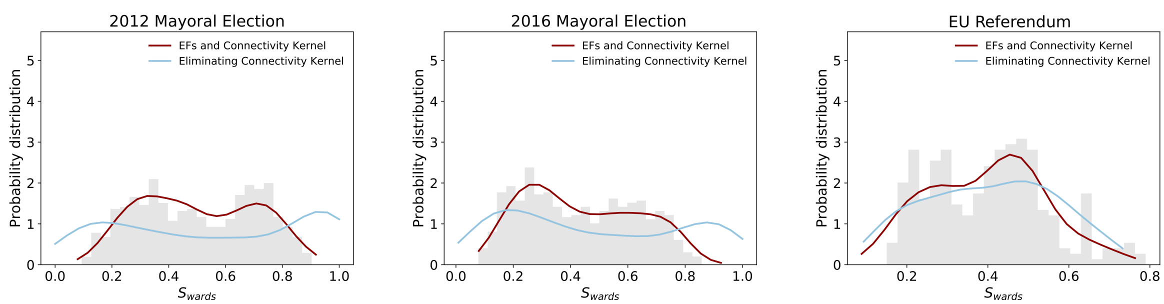

Figure S10: **Electoral outcomes distributions and polarisation**. In the x-axis S_{ward} is the fraction of Conservative/Leave votes in electoral wards, such that for $S_w = 0$ all votes are for Labour/Remain and for $S_w = 1$ all votes are for Conservative/Leave. We show in grey the histograms of the real electoral outcomes; in dark-red solid lines the Gaussian kernel density of the KBI generated outcomes distributions from a multivariate Gaussian regression of the ABC estimates in Fig. 3 in the main text; and in light-blue solid lines the Gaussian kernel density of generated outcomes for the same model parameters but eliminating the connectivity kernels (J=0). We find that for the *eliminating connectivity kernel* ($J = 0$) distributions, the EU Referendum distribution is in better agreement with the original ward outcomes distribution than the corresponding 2012 and 2016 Mayoral Election distributions. This all suggest that for the EU Referendum the kernel adds less new information above the EFs compared to the MEs, as we would expect with outcomes that do not have clearly defined niches in the Blau space.