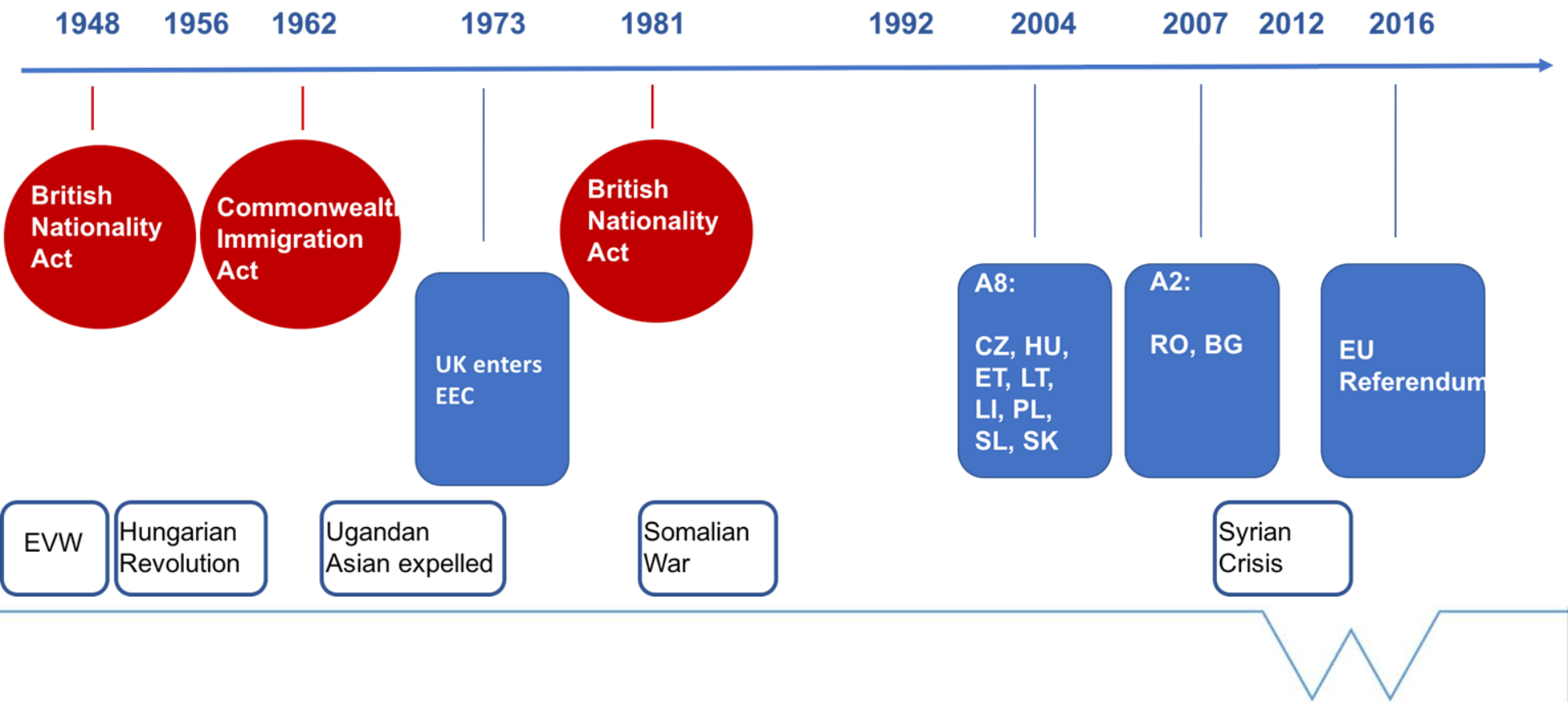
SpaCular - Disclosure of spatial peculiarities of the Brexit

WARWICK

K. Klemmer, J. Porto de Albuquerque, A. Sonea, N. Tkachenko and R. Westerholt // University of Warwick, University of Heidelberg

CDRC Data Challenge - GISRUK 2018 // Leicester // 18/04/2018

UK Immigration - A Timeline



UK Immigration - Context

UK has experienced waves of immigration which have integrated under very different legal frameworks and economic conditions:

- From Commonwealth (the “Windrush” generation, UK citizens under the 1948 British Citizenship Act, later gradually restricted following the 1962 Commonwealth Act)
- From Europe (European Volunteer Scheme after WWII, FOM after the UK accession to ECC in 1973 followed by A8 and A2 extensions in 2004 and respectively 2007)
- Refugees following wars and revolutions (Hungary, 1956; Ugandan Asians, 1972; Somali, 90s; Syria, after 2010)

→ Certain ethnic communities settled in specific areas (e.g largest part of Ugandan community in Leicester)

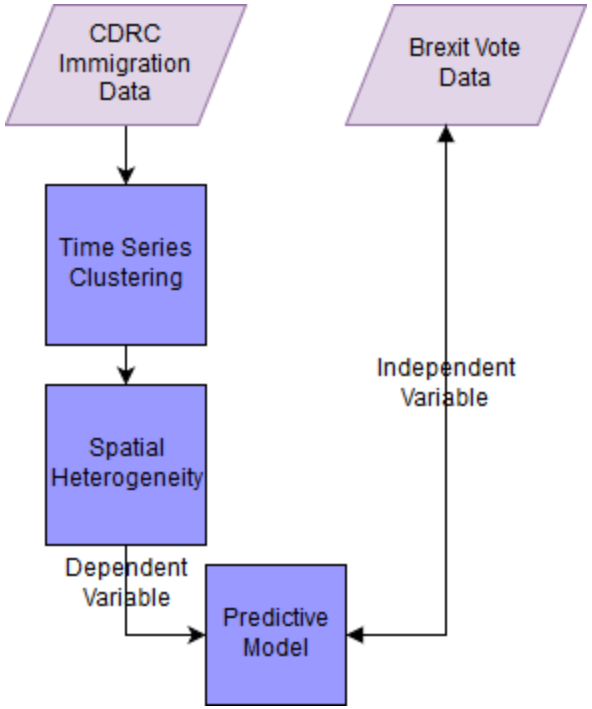
→ Certain areas in UK experienced successive immigration waves since 1948 while others experienced immigration mainly in the past 15 years.

Research Questions

“The change in immigrant population rather then the total number was a driving force in Brexit voting behaviour.”

“Brexit voting behaviour also depends on the immigration profile of neighbouring areas.”

Methodology Overview



Temporal Analysis

Finding **patterns in population development over time** by *ethnicity* and *voting district*

- → Different areas of England exhibit different changes in the population share of different ethnic groups
- → The characteristics of these changes might allow us to group similar areas together
- → This challenge poses an *unsupervised learning problem*
- → **We tackle this issue by applying *time-series clustering***

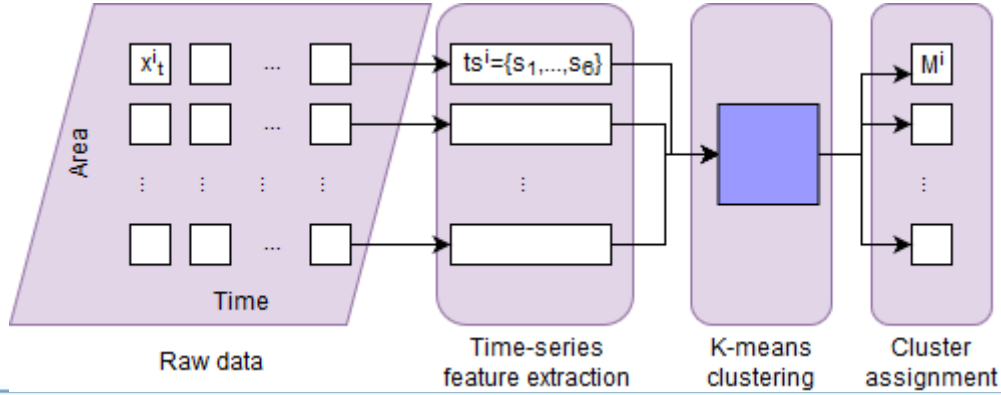
Time-series clustering

Step 1: Extract *features* for each time-series

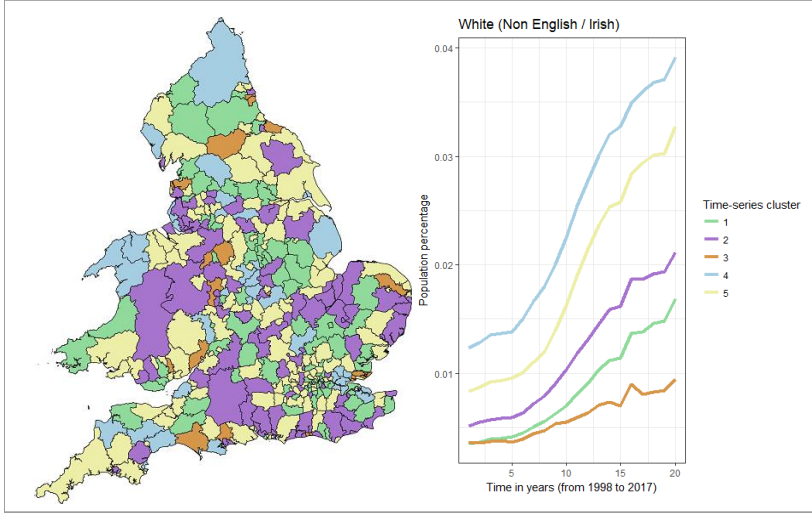
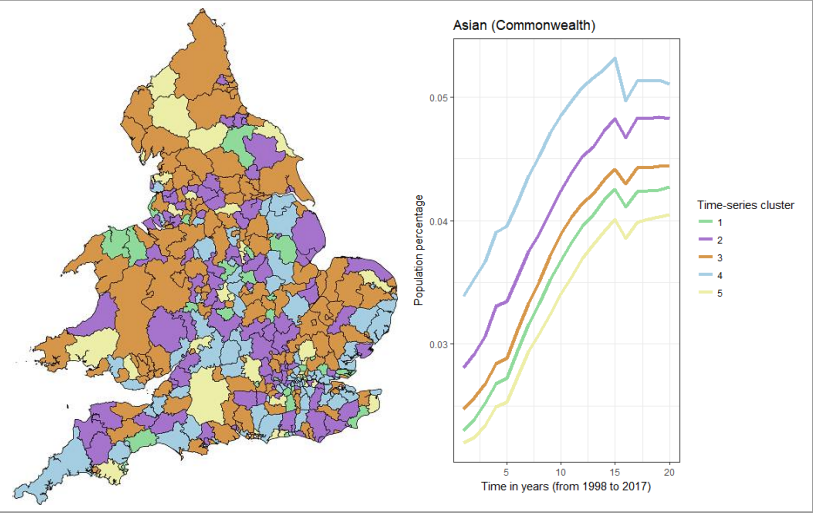
→ *(1)* R^2 and *(2)* estimate of a linear trend model, the *(3)* intercept, *(4)* trend estimate and *(5 and 6)* first two breakpoints of an empirical fluctuation process (EFP).

→ The EFP derives an empirical process for the fluctuation of a given linear (trend) model, hence helps us with extracting information about non-linear time series

Step 2: Apply *k-means clustering* with the extracted time-series features



Time-series clustering

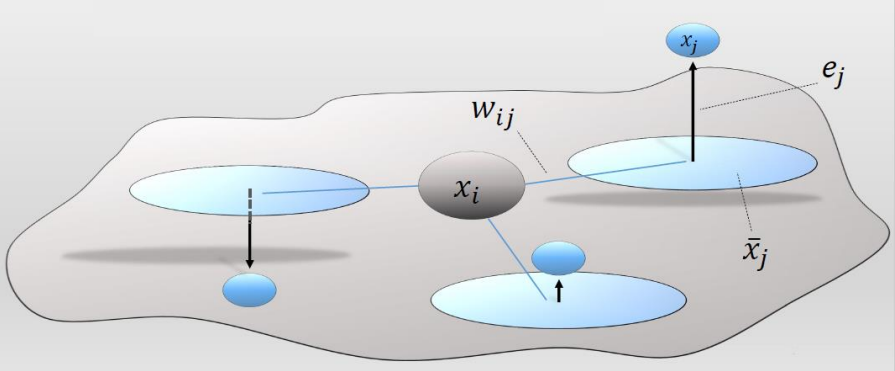
White (Non English / Irish)	Asian (Commonwealth)
	

Spatial Variability (LOcal Spatial Heteroscedasticity LOSH)

Are geographic neighbourhoods more, or less diverse than expected?

→ Estimation of ***spatially weighted variance***

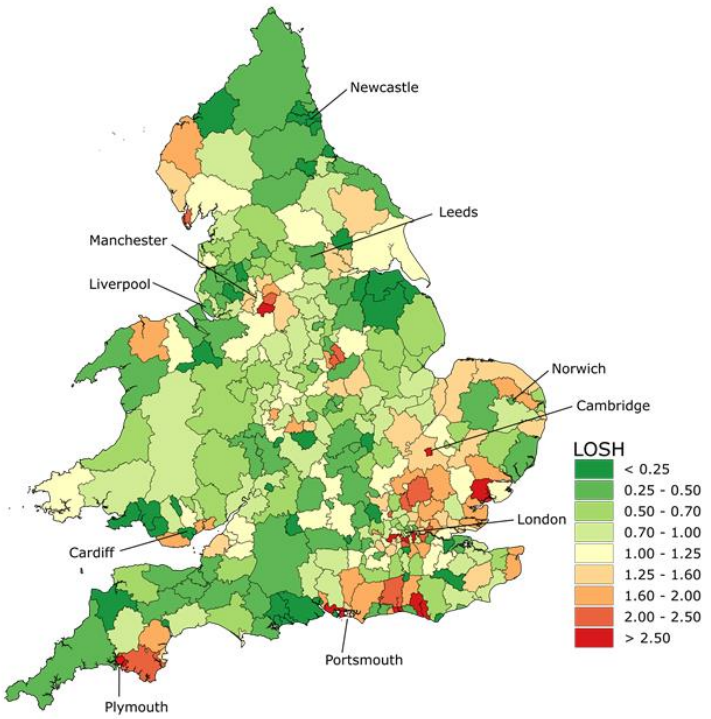
→ Use of a test statistic to disclose ***hot spots***

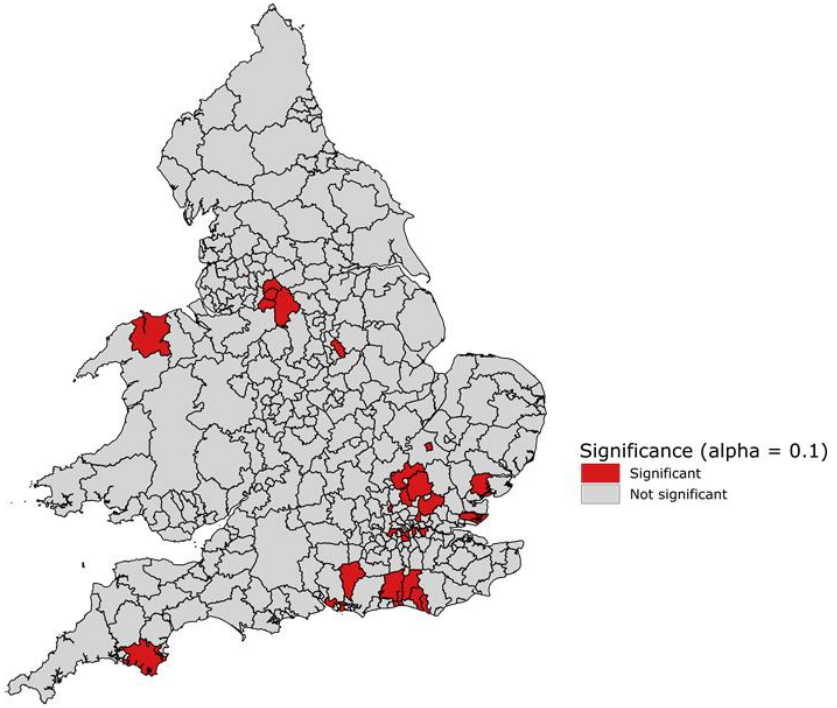


$$H_i = \frac{\sum_j^n w_{ij} \left| e_j \right|^2}{h_1 \sum_j^n w_{ij}}$$

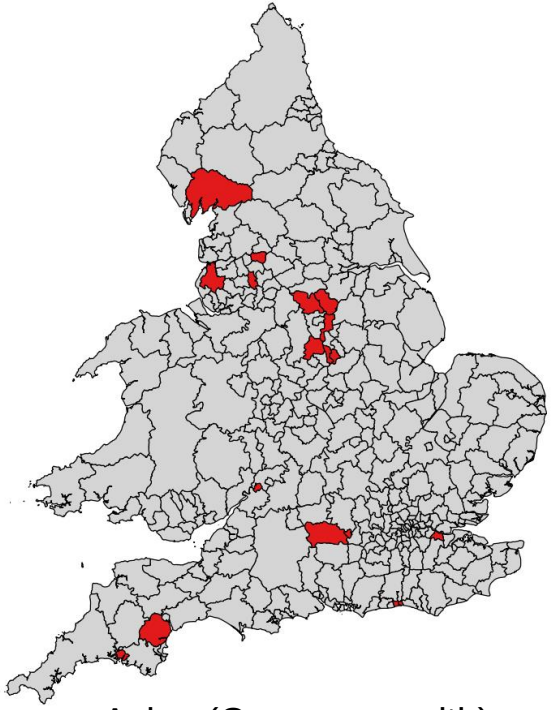
Ord, J. K., & Getis, A. (2012). Local spatial heteroscedasticity (LOSH). *The Annals of Regional Science*, *48*(2), 529-539.

Result: Brexit Vote (% Leave)

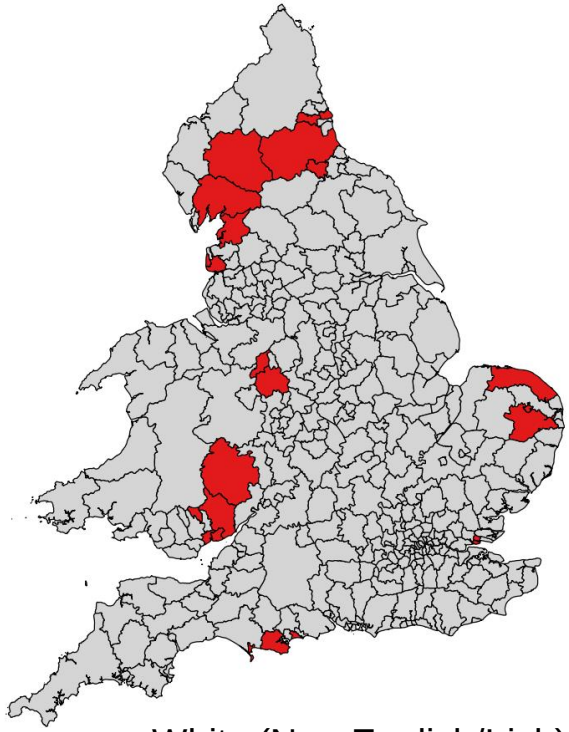




Results: Immigration Dynamics



Asian (Commonwealth)



White (Non-English/Irish)

Significant values
(alpha = 0.05)

Modelling Approach (Dep. var.: Leave %)

1) $\boldsymbol{y} \sim \boldsymbol{x_t}$ ← *Predictor: Ethn. pop. percentage (2017)*
   - **GWR R^2: 0.7, LM R^2: 0.58**
   - Significant predictors: Asian (Chinese), Black (African), Black (Caribbean)
2) $\boldsymbol{y} \sim \boldsymbol{M}(\boldsymbol{x})$ ← *Predictor: Ethn. time-series cluster*
   - GWR R^2: 0.56, LM R^2: 0.12
   - Significant predictors: White (British), White (Other), Black (African), Other
3) $\boldsymbol{y} \sim \Pr(\boldsymbol{H}(\boldsymbol{M}(\boldsymbol{x})))$ ← *Predictor: LOSH p-value of time-series cluster*
   - GWR R^2: 0.51, LM R^2: 0.25
   - Significant predictors: White (Irish), White (Other), Black (African), Black (Other)
4) $\boldsymbol{y} \sim \boldsymbol{x} + \boldsymbol{M}(\boldsymbol{x}) + \Pr(\boldsymbol{H}(\boldsymbol{M}(\boldsymbol{x})))$ ← *Predictor: Combined model*
   - **GWR R^2: 0.8, LM R^2: 0.61**
   - Significant predictors: White (Irish), Asian (Chinese), Black (African), Other

Research Questions

Hypothesis 1 → Hypothesis 2

→ Including information about the (recent) changes in ethn. population enhances the static model.

→ Spatial heterogeneity in ethn. population development plays a substantial role in the Brexit votum.

Space matters!

Outlook

- Apply **spatial cross-validation** to account for overfitting
- Use **dissimilarity-based clustering** for population time series
- **Refine methodology** (e.g. point processes, spatial smoothing)
- Include **socio-economic controls** and insights from previous research

Thank you!

K. Klemmer // k.klemmer@warwick.ac.uk // @kklmmr
J. Porto de Albuquerque, j.porto@warwick.ac.uk // @j_p_albuquerque
N. Tkachenko, n.tkachenko@warwick.ac.uk // @FloodSmartCity
A. Sonea, a.sonea@warwick.ac.uk // @andrasonea
R. Westerholt, westerholt@uni-heidelberg.de // @rwesterh87