Education and the Geography of Brexit

Robert Calvert Jump[1], Jo Michell[2]

[1] *University of the West of England (UWE), Bristol*
[2] *University of the West of England (UWE), Bristol*

Economics Working Paper Series
1901

Education and the geography of Brexit

Robert Calvert Jump* Jo Michell†

July 2, 2019

Abstract

The role of education in the geography of Brexit is usually examined using descriptive statistics and regression, which are ill-suited to the assessment of predictive capacity. By presenting in-sample and out-of-sample probit classification results, this paper demonstrates that educational attainment alone can correctly classify up to 92.24% of local authorities by voting outcome, including up to 80% of Remain-voting authorities. These results emphasise the importance of education as a key factor in the political geography of the Brexit vote.

Keywords: Brexit, Education, Geography.

Acknowledgements: We thank Thiemo Fetzer for comments.

*Department of Accounting, Economics and Finance, University of the West of England. Email: rob.calvertjump@uwe.ac.uk.

†Department of Accounting, Economics and Finance, University of the West of England. Email: jo.michell@uwe.ac.uk.

1 Introduction

In October 2017, the Conservative MP for Daventry wrote to the vice chancellors of British universities to ask for the names of any professors involved in the teaching of European affairs, "with particular reference to Brexit" (Fazackerley, 2017). Discussing this letter on the BBC's Sunday Politics programme, Barry Sheerman, the Labour MP for Huddersfield, argued that, "The truth is that when you look at who voted to Remain, most of them were the better-educated people in our country". This observation provoked widespread debate, with at least one Conservative MP accusing Sheerman of snobbery (Baynes, 2017).

Since this short-lived media storm, and others like it, empirical work has established educational attainment as one of the strongest correlates of voting intention in Britain's referendum on EU membership. This has been found using individual-level data on attitudes towards the EU (e.g Alabrese et al., 2018) and geographic data on voting outcomes (e.g. Becker et al., 2017; Zhang, 2018). The association between education and Brexit is mirrored in the wider literature linking education and Euroscepticism, including Lubbers & Scheepers (2010) and Hakhverdian et al. (2013).

While Alabrese et al. (2018) conducts a classification exercise using individual-level data, the existing literature on the geography of Brexit concentrates on descriptive statistics and regression. In doing so, these studies miss the remarkable ability of educational attainment to predict the geographic distribution of the Leave vote. The present paper fills this gap in the literature, by using single and multi-level probit models to classify local authorities in England and Wales into Leave and Remain areas using a variety of demographic predictors. Both the in-sample and out-of-sample classification success of educational attainment is remarkable, and emphasises its importance as a key predictor of Britain's vote to leave the European Union.

2 Data and descriptive statistics

The data are comprised of Leave vote shares and a set of demographic characteristics, observed at local authority level in England and Wales. The latter comprise the population shares born in the UK, identifying as ethnically white, identifying as male, median age, and the working population share with a "lower supervisory and technical", "semi-routine", or "routine" occupation. These are all standard predictors in the existing literature. The measure of educational attainment is the adult population share educated to degree-level or above, which is 27% on average across English and Welsh local authorities. Voting data are taken from the Electoral Commission and the demographic variables are taken from the 2011 census. We confine our analysis to England and Wales, because data for Northern Ireland are unavailable at local authority level and every authority in Scotland voted Remain.

3 Classification Approach

To classify local authorities into Leave and Remain voting areas, we estimate probit models of the form,

$$\Pr[y_{ij} = 1|\mathbf{x}_{ij}, u_j] = \Phi(\mathbf{x}_{ij}\beta + u_j), \tag{1}$$

for local authorities indexed by i in NUTS 1 regions indexed by j, where Φ is the c.d.f. of the standard Normal distribution, $\mathbf{x}_{ij}$ are the set of demographic regressors, and u_j are a set of fixed or random regional effects. The dependent variable $y_{ij} = 1$ if the local authority voted Leave, and 0 otherwise. The models in (1) yield a set of predicted probabilities $\hat{\Pr}$. If we define a correctly classified Leave area as one for which $y = 1$ and $\hat{\Pr} > 0.5$, and a correctly classified Remain area as one for which $y = 0$ and $\hat{\Pr} \leq 0.5$, then we can calculate the percentage of correctly classified Leave areas, the percentage of correctly classified Remain areas, and the percentage of correctly classified areas overall. Taken together, these statistics provide a useful summary of predictive accuracy.

4 Results

Table 1 presents in-sample classification results using standard, fixed effect, and random effect models. The fixed effects and random effects model perform similarly, and both are more successful than the standard probit models. Remarkably, the fixed effects model with education as its sole predictor correctly classifies 96.2% of Leave-voting areas and 80% of Remain-voting areas. In absolute numbers, the model correctly classifies 253 of the 263 Leave-voting areas and 68 of the 85 Remain-voting areas. In comparison, the fixed effects model using all of the demographic variables correctly classifies 96.96% of Leave-voting areas and 82.35% of Remain-voting areas, or 255 and 70 correctly classified Leave and Remain areas respectively. As such, the marginal gain of using the full set of demography variables over using education as the sole predictor is two correctly classified Leave areas and two correctly classified Remain areas.

In the machine learning literature, the validation of classification models usually involves estimation on a training set followed by classification exercises performed on a validation set, in which observations are randomly assigned between sets. This is the approach taken in Alabrese et al. (2018), for example. However, because we have only 348 observations, taking this approach is likely to result in considerable sampling variability in any measure of predictive success. Instead, table 1 presents classification results for local authorities in the South East predicted by standard probit models estimated on the remaining NUTS 1 regions. We use the South East for this exercise because its Leave share (51.78%) was very similar to the overall result (51.89%), and it has the largest electorate of the NUTS 1 regions.

Table 1: Classification results

	Correctly classified districts:		
	Overall	Leave	Remain
Standard probit in-sample:			
Education only	90.80%	95.44%	76.47%
All demography	92.24%	96.96%	77.65%
All demography minus education	85.92%	93.54%	62.35%
Fixed effects probit in-sample:			
Education only	92.24%	96.20%	80.00%
All demography	93.39%	96.96%	82.35%
All demography minus education	89.94%	95.44%	72.94%
Random effects probit in-sample:			
Education only	92.24%	96.96%	77.65%
All demography	93.10%	97.34%	80.00%
All demography minus education	88.79%	95.44%	68.24%
Standard probit out-of-sample:			
Education only	91.04%	88.37%	95.83%
All demography	88.06%	90.70%	83.33%
All demography minus education	80.60%	81.40%	79.17%

Notes: The standard probit models do not include fixed or random regional effects. The random effects models are estimated using `meprobit` in Stata. The models for the in-sample results are estimated using all local authorities England and Wales. The models for the out-of-sample results are estimated on all NUTS 1 regions other than the South East, which is used for the out-of-sample classification exercise.

The model performs relatively well in comparison to the in-sample classifications, correctly classifying 38 of the 43 Leave districts and 23 of the 24 Remain districts. Interestingly, the model with education as its sole predictor outperforms the model with every demographic variable, which is comparatively poor at classifying the Remain areas.

5 Discussion

The local authorities incorrectly classified by the random effects probit model are displayed in figure 1. Misclassified Remain-voting areas include several large cities and towns with large universities, suggesting a possible influence of education that is not captured by the proportion of degree-educated adults. On the other hand, most of the other incorrectly classified Remain areas are geographically contiguous with areas with particularly strong Remain votes. There is, therefore, clearly information being ignored by our models. Nonetheless,

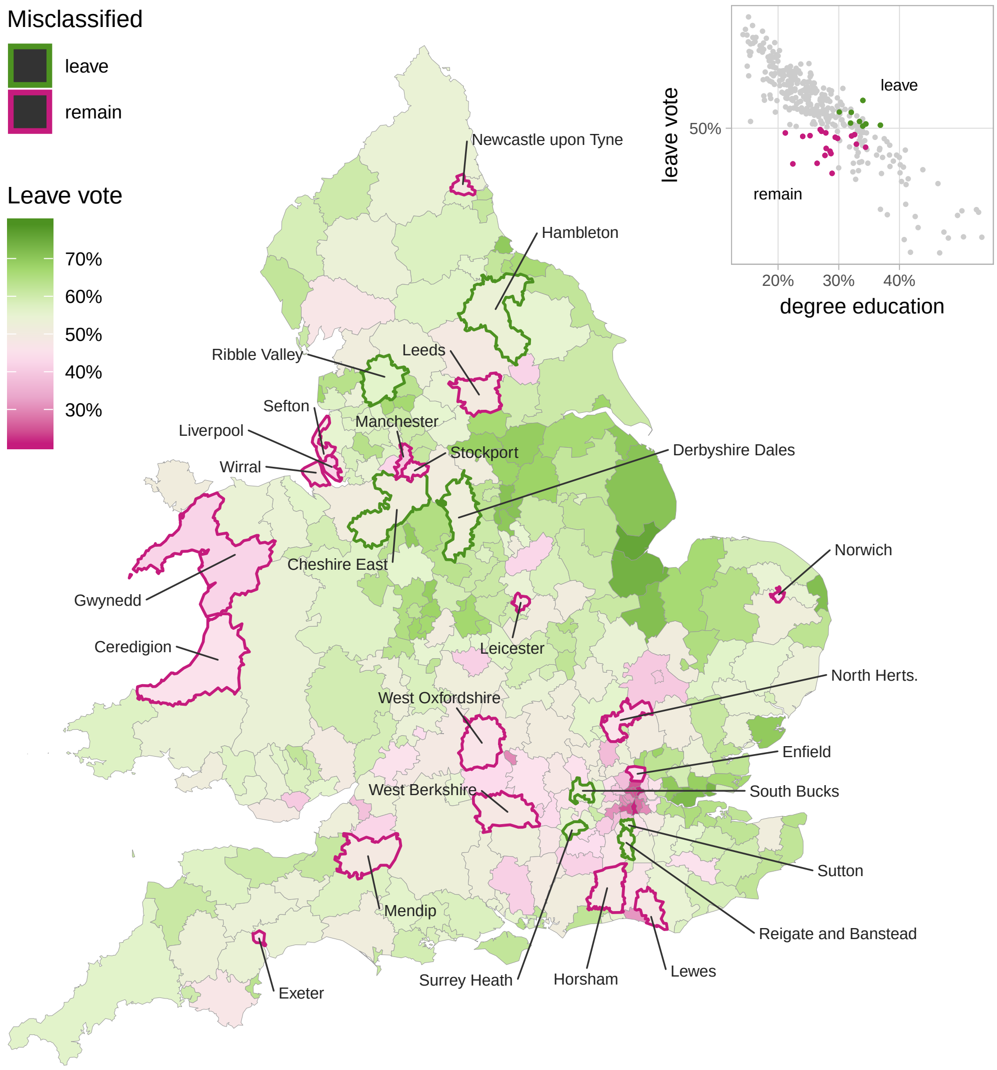

Figure 1: Map of Leave vote share, with incorrectly classified areas highlighted. Inset scatter shows Leave vote share against degree education.

probit models using only education have remarkable predictive capacity for the geographic distribution of Britain's vote to leave the EU.

How should this result be interpreted in light of the various narratives seeking explain the Brexit vote? These can broadly be divided into those emphasising cultural divergence between socially liberal Remain voters and socially conservative Leave voters (e.g. Goodwin & Heath, 2016), and those emphasising economic drivers such as inequality, austerity, and the effects of globalisation (e.g. Fetzer, 2018). Since educational attainment is closely correlated with both social attitudes and economic success, our results could be invoked in favour of either set of narratives. We therefore caution against drawing firm conclusions on the basis of these results. Instead, the predictive capacity of education is one clue in the larger puzzle of what drove the Brexit vote.

References

Alabrese, E., Becker, S., Fetzer, T., & Novy, D. (2018). Who voted for brexit? individual and regional data combined. *European Journal of Political Economy*.

Baynes, C. (2017). Brexit: Labour mp who said remain voters were better educated 'has facts on his side', says former yougov boss. *The Independent*.
URL https://www.independent.co.uk/news/uk/politics/brexit-remain-voters-better-educated-labour-mp-barry-sheerman-huddersfield-yougov-facts-leave-a8027121.html

Becker, S., Fetzer, R., & Novy, D. (2017). Who voted for brexit? a comprehensive district-level analysis. *Economic Policy*, *32*(92), 601–650.

Fazackerley, A. (2017). Universities deplore 'mccarthyism' as mp demands list of tutors lecturing on brexit. *The Guardian*.
URL https://www.theguardian.com/education/2017/oct/24/universities-mccarthyism-mp-demands-list-brexit-chris-heaton-harris

Fetzer, T. (2018). Did austerity cause Brexit? Working paper, University of Warwick.
URL http://wrap.warwick.ac.uk/106313/

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

Hakhverdian, A., Van Elsas, E., Van der Brug, W., & Kuhn, T. (2013). Euroscepticism and education: A longitudinal study of 12 eu member states, 1973–2010. *European Union Politics*, *14*(4), 522–541.

Lubbers, M., & Scheepers, P. (2010). Divergent trends of euroscepticism in countries and regions of the european union. *European Journal of Political Research*, *49*(6), 787–817.

Zhang, A. (2018). New findings on key factors influencing the uk's referendum on leaving the eu. *World Development*, *102*, 304–314.

Recent UWE Economics Papers
See https://www1.uwe.ac.uk/bl/research/bcef/publications.aspx for a full list.

2018

1807 **Learning, Heterogeneity, and Complexity id the New Keynesian Model**
Robert Calvert Jump, Cars Homme, and Paul Levine

1806 **DSGE Models and the Lucas Critique. A Historical Appraisal**
Francesco Sergi

1805 **A new approach to estimating interregional output multipliers using input-output data for South Korean Regions**
Malte Jahn, Anthony T. Flegg and Timo Tohmo

1804 **Urban food security in the context of inequality and dietary change: a study of school children in Accra**
Sara Stevano, Deborah Johnston and Emmanuel Codjoe

1803 **The use of differential weighting and discounting in degree algorithms and their impact on classification inflation and equity: A further analysis**
David O. Allen

1802 **Unambiguous inference in sign-restricted VAR models**
Robert Calvert Jump

1801 **Degree algorithms, grade inflation and equity: the UK higher education sector**
David O. Allen

2017

1706 **Internal rationality, heterogeneity and complexity in the new Keynesian model**
Cars Hommes, Robert Calvert Jump and Paul Levine

1705 **The regionalization of national input–output tables: a study of South Korean regions**
Anthony T. Flegg and Timo Tohmo

1704 **The impact of quantitative easing on aggregate mutual fund flows in the UK**
Iris Biefang-Frisancho Mariscal

1703 **Where are the female CFOs?**
Gail Webber, Don J Webber, Dominic Page and Tim Hinks

1702 **Mental health and employment transitions: a slippery slope**
Don J Webber, Dominic Page and Michail Veliziotis

1701 **SMEs access to formal finance in post-communist economies: do institutional structure and political connectedness matter?**

Kobil Ruziev and Don J Webber

2016

1611 **Curriculum reform in UK economics: a critique**
Andrew Mearman, Sebastian Berger and Danielle Guizzo

1610 **Can indeterminacy and self-fulfilling expectations help explain international business cycles?**
Stephen McKnight and Laura Povoledo

1609 **Pricing behaviour and the role of trade openness in the transmission of monetary shocks**
Laura Povoledo

1608 **Measuring compliance with minimum wages**
Felix Ritchie, Michail Veliziotis, Hilary Drew and Damian Whittard

1607 **Can a change in attitudes improve effective access to administrative data for research?**
Felix Ritchie

1606 **Application of ethical concerns for the natural environment into business design: a novel business model framework**
Peter Bradley, Glenn Parry and Nicholas O'Regan

1605 **Refining the application of the FLQ Formula for estimating regional input coefficients: an empirical study for South Korean regions**
Anthony T. Flegg and Timo Tohmo

1604 **Higher education in Uzbekistan: reforms and the changing landscape since independence**
Kobil Ruziev and Davron Rustamov

1603 **Circular economy**
Peter Bradley

1602 **Do shadow banks create money? 'Financialisation' and the monetary circuit**
Jo Michell

1601 **Five Safes: designing data access for research**
Tanvi Desai, Felix Ritchie and Richard Welpton

2015

1509 **Debt cycles, instability and fiscal rules: a Godley-Minsky model**
Yannis Dafermos