Who Voted for Brexit?
A Comprehensive District-Level Analysis

Sascha O. Becker Thiemo Fetzer Dennis Novy*

June 2017

Abstract

On 23 June 2016, the British electorate voted to leave the European Union. We analyze vote and turnout shares across 380 local authority areas in the United Kingdom. We find that exposure to the EU in terms of immigration and trade provides relatively little explanatory power for the referendum vote. Instead, we find that fundamental characteristics of the voting population were key drivers of the Vote Leave share, in particular their education profiles, their historical dependence on manufacturing employment as well as low income and high unemployment. At the much finer level of wards within cities, we find that areas with deprivation in terms of education, income and employment were more likely to vote Leave. Our results indicate that a higher turnout of younger voters, who were more likely to vote Remain, would not have overturned the referendum result. We also compare our UK results to explain the vote shares of the far-right leader Marine Le Pen in the 2017 French presidential election. We find similar factors driving the French vote. An out-of-sample prediction of the French vote using UK estimates performs reasonably well.

Keywords: Political Economy, Voting, Referendum, Migration, Austerity, Globalisation, UK, Scotland, EU, France
JEL Classification: D72, N44, R23, Z13

Acknowledgements: Paper presented at the 65th Panel Meeting of *Economic Policy* in April 2017. We are grateful for constructive comments received by the panel, our discussants Vasso Ioannidou and Andrea Mattozzi, the editor and three referees. We also thank Andrew Oswald, Alan Taylor, seminar participants at Warwick and HM Treasury as well as numerous commentators on social media. We thank the ESRC initiative "The UK in a Changing Europe" for financial support. We thank Bart Los for sharing data with us. Eleonora Alabrese and Ivan Yotzov provided outstanding research assistance. All three authors are based at the University of Warwick and affiliated with the Centre for Competitive Advantage in the Global Economy (CAGE). Becker is also affiliated with CEPR, CESifo, ifo, IZA and ROA. Fetzer is also affiliated with SERC. Novy is also affiliated with CEPR, CESifo and CEP/LSE. *Email addresses*: Becker: s.o.becker@warwick.ac.uk; Fetzer: t.fetzer@warwick.ac.uk; Novy: d.novy@warwick.ac.uk.

1 Introduction

The United Kingdom's relationship with the European Union (EU) has always been a very special one. Not being a founding member, the UK only joined the European Economic Community (EEC), the precursor of the EU, in 1973. Merely two years later, the UK held its first in-out referendum. It produced a clear two-thirds majority to remain as a member. The UK has historically been a key supporter of several core features of today's EU such as the Single Market and EU Regional Policy. However, the UK never joined the Euro. It did not follow the route of the six founding members of the European project (and many other countries) to proceed towards ever closer union (see Guiso et al., 2016 for an analysis of Euro membership in the context of the European integration project).[1] Over the last two decades, the UK seemed to have grown increasingly lukewarm towards the EU. During the 2015 general election campaign, internal struggles within David Cameron's Conservative party led him to promise a referendum on EU membership. This referendum happened on 23 June 2016.

The UK referendum on EU membership is thought to have been a watershed moment in European integration and globalisation more broadly. Although the outcome had been expected to be tight, in the days running up to the referendum bookmakers and pollsters predicted the Remain side to win. Many observers were left puzzled and keen to understand who voted for Leave. Various newspapers and blogs quickly produced correlations between selected variables and the referendum result, but no study has so far taken a comprehensive approach to attempting to understand the Brexit vote.[2] Our paper fills this gap by combining a multitude of geographically disaggregated data sources to carry out a comprehensive descriptive analysis of the socio-economic characteristics that correlate with the outcome of the 2016 referendum.

In particular, we study the EU referendum results in England, Wales and Scotland disaggregated across 'local authority areas' of the referendum (and across 107 wards within four cities) and relate them to fundamental socio-economic features of these areas.[3] The EU Referendum Act passed by Parliament in 2015 divided the UK into 382 official counting areas (which are the same as local authority areas), 327 of which are in

[1] See Appendix A for a more detailed history of Britain's role in the EU.

[2] For instance, see Burn-Murdoch (2016b) in the Financial Times as an example of various correlation plots.

[3] An analysis of voting at the local authority area level does not necessarily reflect individual voting behaviour, a phenomenon called 'ecological fallacy'. We deliberately want to understand these regional voting patterns. But our analysis of within-city variation goes one step towards addressing worries about ecological fallacy because of the much finer level of geographical disaggregation.

England, 22 in Wales and 32 in Scotland.[4] There are on average roughly 122,000 eligible voters per local authority area. Data is not provided at the level of individual polling stations.

As covariates, we focus on socio-economic characteristics that can be broadly grouped into four categories: measures of an area's exposure to the European Union; measures capturing (the quality of) public services provision and exposure to fiscal consolidation (austerity); demographic and human capital characteristics; and measures capturing the underlying economic structure of an area.

We adopt a simple machine learning method to capture the subsets of variables from each group that best 'predict' the actual referendum result. We cannot possibly give a causal explanation of the referendum result because the election outcome is multi-causal and multi-faceted. Nevertheless, a systematic analysis across an exhaustive range of socio-economic characteristics can be helpful in directing future research efforts that aim at carefully identifying specific mechanisms. One might be able to single out an individual predictor such as immigration from Eastern Europe and try to establish causality from this specific factor for the Vote Leave share. But this would run counter to the aim of this paper, which is rather to focus on predictive power by pulling together various dimensions of the vote pattern.

Our results indicate that even very simple empirical models can explain significant amounts of variation in the Vote Leave share and achieve good predictive performance. Which characteristics have significant explanatory power for Leave support? Surprisingly and contrary to much of the political debate in the run-up to the election, we find that relatively little variation in the Vote Leave share is explained by measures of a local authority area's exposure to the European Union (e.g., due to immigration and trade exposure). Neither is much variation explained by measures capturing the quality of public services and fiscal consolidation. Rather, a significant amount of the variation can be linked to variables that seem hardly malleable in the short run by political choices (variables such as educational attainment, demography and industry structure). We document that similar patterns hold when we explore data on the EU referendum result across 107 wards in four English cities – which to the best of our knowledge this paper is the first to exploit.

Our findings thus suggest that there is a disconnect between the key correlates of

[4]We drop Northern Ireland because election results were only published for Northern Ireland as a whole. This makes Northern Ireland an outlier by being the largest "local authority" by an order of magnitude. We also drop Gibraltar, a British Overseas Territory ceded to Britain in 1713 under the Treaty of Utrecht, where many covariates are missing. Thus, we end up with 380 voting observations at the local authority level.

the vote outcome and the topics dominating the political debate in the run-up to the election. How can we reconcile this disconnect? The political debate centred on two issues: the fiscal burden of EU membership and the exposure to European immigration since the enlargement of the European Union in 2004. Perhaps the UK budget contribution resonated so strongly with the British electorate because public services and benefits were under severe strain not least due to fiscal cuts. If we think of fiscal cuts and migration as political choice variables, we can explore the extent to which the powerful predictors capturing the underlying fundamentals (educational attainment, demography and industry structure) interact with these variables that saw significant change over the course of the last decade. Our results highlight that policy choices related to pressure from fiscal cuts and migration are linked to a higher Vote Leave share especially when socio-economic fundamentals are 'weak' (low incomes, high unemployment), and when the local population is less able to adapt to adverse shocks (due to low qualifications).

We stress that whilst our paper focuses on the *variation* of vote shares across local authority areas with respect to key variables such as immigration and education, we have less to say about the overall *level* of support for Vote Leave. Put differently, our paper focuses on slope coefficients, not intercepts. This is important because in order to get a sense of the absolute number of people who voted for or against Brexit, one would need to refer to data on individuals and how they voted. To some extent, such information is available through polling data, for instance as provided by Ashcroft (2016). Such polls indicate that the typical Leave voter is white, middle class and lives in the South of England. The proportion of Leave voters that are in the lowest two social classes (D and E) is less than one-third (see Dorling, 2016).

We also carry out a back-of-the-envelope calculation regarding turnout. Young people voted overwhelmingly in favour of Remain but had a lower turnout than older age groups. We find that a higher turnout of young voters would have been very unlikely to result in a different referendum outcome, partly because their turnout was already elevated compared to previous UK-wide elections.

We also explore the role of some short-run factors such as heavy rainfall and flooding on the referendum day as well as train cancellations in the South East of England. While we document that these did have a reducing effect on turnout, the reduction does not seem to have affected the overall result: the Remain campaign would have still lost on a sunny day.

Lastly, we also compare our UK results to explain the vote shares of the far-right

leader Marine Le Pen across départements in the 2017 French presidential election. Arguably, both the Leave vote and the support for Le Pen can be described as having a distinct populist flavour. The question is whether both votes are related to similar underlying socio-economic conditions. We find that the factors driving the French vote are indeed similar to those in the UK. A corresponding model for France using the same variables as for the UK has explanatory power not far below that for the Brexit referendum. Even an out-of-sample prediction of the French vote using UK estimates performs reasonably well.

This paper and the Brexit vote it studies can be seen not only in an EU context but also related to 'populist' campaigning and voting more broadly. A large literature in the social sciences looks at voting patterns across the political spectrum as a function of demographic, economic and political drivers (see Ferree et al., 2014). The UK, with its first-past-the-post electoral system for the House of Commons, has typically had clear majorities for either the Conservatives or the Labour Party since the 1920s. This pattern was broken in 2010 with the first coalition government that saw the Conservatives and the Liberal Democrats join forces. Since the 1990s two other major developments have affected the UK party landscape: the rise of the UK Independence Party (UKIP) and the rise of 'nationalist' parties in Scotland, Wales and Northern Ireland. While the latter can be seen as a domestic move driven by a renewed push for devolution (and even independence) for the constituting nations of the UK, the rise of UKIP is directly related to the EU. Whitaker and Lynch (2011) as well as Clarke et al. (2016) study voting patterns for UKIP and document that, not surprisingly, Euroscepticism combined with anti-immigration sentiment is the main driving force of UKIP success. For Western Europe more broadly, Arzheimer (2009) analyzes contextual factors explaining far-right voting over the period from 1980 to 2002.

Backlash against globalisation is said to have been another important factor in the Leave vote, especially to the extent that it deteriorates economic and social conditions for a subset of voters (see Lewis-Beck and Stegmaier, 2000 and Druckman and Lupia, 2000).[5] Colantone and Stanig (2016) provide evidence of import competition from China being related to support for Vote Leave in an arguably causal manner. Their results are consistent with ours in two ways. First, we also find a positive relationship between trade intensity (in our case with other EU countries) and support for Vote Leave. Second, we confirm that areas heavily dependent on manufacturing employment were

[5] Findlay and O'Rourke (2007) argue that globalisation was historically difficult to maintain unless domestic institutions developed and adapted accordingly. This often meant a strong role for the state, for instance in the form of educational, training and welfare programmes.

more likely to vote Leave.

Of course, the UK's Brexit vote should not be equated with support for UKIP or far-right voting more generally. Yet, there are probably some parallels with voting patterns for right-wing parties in other countries and the 'once-in-a-lifetime' opportunity to vote against what many voters see as unaccountable forces ruling them from the outside. In the UK context, Becker and Fetzer (2016) explore the impact of immigration from Eastern Europe on the support for UKIP. Dippel et al. (2015) link votes for far-right parties in Germany to trade integration with China and Eastern Europe. For the U.S., Autor et al. (2016) argue that rising trade integration with China contributes to the polarization of U.S. politics. Burgoon (2012) analyzes party opposition and support for trade openness across the European Union. Barone et al. (2016) find that in Italy, immigration generates a sizeable causal increase in votes for the centre-right coalition that has a political platform less favourable to immigrants.

The UK's EU referendum is of course also related to research on referenda as a form of direct voting. While countries such as Switzerland have ample experience in 'direct voting' (see Funk and Gathmann, 2015), referenda in other countries are rather rare. The UK traditionally respects the primacy of Parliament over any direct voting. But both the UK's European Economic Community (EEC) referendum in 1975 and the EU referendum in 2016 were initiated by the House of Commons. Theoretical research has come up with suggestions to improve the efficiency of referenda (Casella and Gelman, 2008). On the empirical side, Matsusaka (1992) asks why some issues are resolved by popular vote and others by elected representatives. Using data on California he finds that "good government" issues were usually resolved by legislative measures and distributional issues by initiatives. In light of this finding, it makes sense to view the Brexit referendum as one that was at least partially related to distributional issues.

Our paper is organised as follows. Section 2 introduces our empirical approach. Section 3 discusses the underlying data and our main hypotheses. In Section 4 we present our results. Section 5 provides a summary and policy conclusions.

2 Empirical approach

We take a comprehensive approach to understanding the factors behind the EU referendum result, and we exploit a range of data sources in the empirical analysis. We would like to stress right away that our analysis cannot possibly establish causality. Instead, we try to capture predictive power of various groups of regressors to see which factors

explain a larger share of the variation in the Vote Leave share. This approach is quite natural in this setting with a once-in-a-lifetime referendum where we are bound to analyze cross-sectional variation only. If we were to analyze general election results, we could recur to difference-in-difference type estimates in order to control for fixed effects at the local authority level. In our analysis, we do not necessarily expect coefficient signs of each and every coefficient to be stable across all specifications. Instead, it is expected that the signs of some regressors, to the extent that they are highly correlated with each other, may change when "more fundamental" regressors are added. We will discuss all these issues in more detail when interpreting our results.

We carry out three main exercises: a full model, a best subset selection procedure, and a within-city analysis. We describe these here in turn. Readers familiar with model selection procedures may want to jump to section 2.3.

2.1 Full model

The first approach aims at building a 'full' empirical model of the correlation structure between k-dimensional cross-sectional covariates X_c at the local authority area level (380 spatial units across England, Scotland and Wales) and a dependent variable y_c, which is either the share of votes to leave L_c or turnout T_c.[6]

For time-varying observables, the cross-sectional covariates contain their respective baseline levels (mostly from the 2001 census), x_{ct}, as well as their changes, Δx_c, mostly between 2001 and 2011, the two census years. The empirical specification takes the form

$$y_c = \mathbf{x}_c'\beta + \epsilon_c, \tag{1}$$

which we estimate with ordinary least squares (OLS).

2.2 Model selection

In the second approach, we perform a variable selection exercise to identify the most robust predictors of the Vote Leave result. In order to identify robust predictors of the Vote Leave result, we perform a best subset selection (BSS) procedure. Best subset selection is a machine learning method used to perform 'feature selection' in settings where the aim is to reduce dimensionality of a feature space (Guyon and Elisseeff,

[6]We remind readers that we drop Northern Ireland and Gibraltar. Northern Ireland is dropped because referendum results were only published for the whole region, at a much more aggregated level than for all other parts of the UK. Gibraltar, a British Overseas Territory, has many missing covariates. Our results are robust to the inclusion of these additional observations.

2003). The idea of best subset selection is to estimate all possible regressions including all combinations of control variables and return the statistically optimal model, which minimizes an information criterion.

The fundamental difference between prediction, which generally takes advantage of machine learning methods, and causal inference is as follows. While causal inference focuses on the internal validity of causally estimated reduced-form (or structural) parameters β, prediction and thus machine learning is concerned with the external validity of the estimated fitted values $\hat{y}$. Causal inference seeks to obtain a set of estimated parameters $\hat{\beta}$ that are typically studied in isolation. Thus they often do not render themselves useful for predictive exercises since the out-of-sample model fit is generally poor. Instead, good model fit typically requires a multitude of regressors, and machine learning can often substantially improve out-of-sample predictive performance (Mullainathan and Spiess, 2017).[7] The underlying estimated parameters that yield good model fit are typically of limited interest per se.[8]

We note that the variables we consider pass a first plausibility test (as they were mentioned during the campaign, for example). They cover broad socio-economic characteristics. They are related to the political science literature documenting determinants behind elections (we refer to that literature in the introduction). They do not contain 'nonsensical' variables that could be thought of as generating 'random' and thus meaningless correlations.

The best subset selection algorithm we employ finds the solution to the following non-convex combinatorial optimization problem:

$$min_{\beta} \underbrace{\sum_{c=1}^{C} (y_c - \beta_0 - \sum_{j=1}^{p} x_{cj}\beta_j)^2}_{\text{Residual sum of squares}} \text{ subject to } \sum_{j=1}^{p} \mathbf{I}(\beta_j \neq 0) \leq s, \tag{2}$$

where p is the set of regressors of which a subset s is chosen to maximize overall model fit. The result is a sequence of models $\mathcal{M}_1, ..., \mathcal{M}_s, .., \mathcal{M}_p$, where the overall optimal model $\mathcal{M}_{s*}$ is chosen by using either cross validation or some degree-of-freedom-adjusted measure of goodness of fit such as the Akaike information criterion (AIC). Throughout, we use the AIC to decide upon the overall optimal model $\mathcal{M}_{s*}$ robustly explaining the variation in the dependent variable.

It is easy to see that this statistically optimal procedure can quickly become infeasi-

[7] See section 4.5 where we predict out-of-sample the results of the 2017 French presidential election.

[8] Some machine learning methods are non-parametric to the extent that the methods do not even produce any model parameters in a classical regression sense.

ble. Suppose there are p potential regressors. Best subset selection proceeds as follows: the first model estimates – using OLS – all $\binom{p}{1} = p$ different models containing a single regressor and chooses as optimal the model that results in the largest reduction in the residual sum of squares. The second model estimates all possible $\binom{p}{2}$ models containing exactly two regressors, and so on. In total, $\sum_{k=1}^{p} \binom{p}{k} = 2^p$ models are estimated. With $p = 30$ this amounts to estimating just over one billion regressions. The non-feasibility of best subset selection for large p in high dimensional data has led to machine learning research efforts focusing on developing algorithms that solve an approximation of the best subset selection optimization problem such as Lasso, Ridge regression or Forward/Backward stepwise selection (see Hastie et al., 2009 for an overview).

It is important to highlight that the best subset selection approach may yield models of different complexity that are non-nested. We present the sequence of "best" models for each class of models with p predictors and explore how the inclusion of more covariates expands the goodness of fit. One caveat with this approach is that certain variables may be dropped in case they are highly correlated with each other. That is, even if a predictor x_i contains a distinct signal conditional on x_j, it may be dropped from the analysis as the signal contained is not sufficiently strong.

2.3 Within-city analysis

While official results are only published at the level of local authority areas, we also managed to obtain voting data at the ward level across four UK cities (see section C.1 in Appendix C for a description). This allows us to zoom into city wards. It also allows us to address potential worries about ecological fallacy. There is ample variation in the Vote Leave shares within cities. As a matter of fact, the variation within cities is larger than across local authorities.

2.4 (No) Difference-in-differences

We considered using the 1975 EU referendum in a difference-in-differences framework. Unfortunately, corresponding data for the 1975 referendum were only published for 68 counting areas across the UK (see Figure A1 for a map of the Leave vote in the 1975 referendum). More importantly, the 1975 referendum took place in a completely different environment. At the time, the Labour party had pledged to hold a referendum. Margaret Thatcher, the newly elected leader of the Conservatives at the time, campaigned for Remain. Remain won with a smashing 67.2 percent vote share. Against this backdrop, a difference-in-difference analysis is not possible. Note, however, that we include

the 1975 referendum vote shares as a regressor in our analysis and generally find a negative correlation between the 1975 Leave share and the 2016 Leave share. This finding attests to the notion that these referenda took place under very different circumstances (see Butler and Kitzinger, 1976 and Crafts, 2016 for further background).

3 Hypotheses and data

In this section we discuss prominent hypotheses that have been proposed to explain the EU referendum result and how we try to capture them in our empirical analysis. We briefly discuss the variables employed in the analysis.

The empirical analysis of UK election data is challenging as the data is provided only at the relatively coarse geographic resolution of 380 local authority areas.[9] We start out in section 3.1 by discussing our main outcome variable, the Vote Leave share in the 2016 referendum, as well as turnout and then turn to the explanatory factors behind the outcome. For these factors, we will look at four broad groups of variables:

1) EU exposure through immigration, trade and structural funds;

2) local public service provision and fiscal consolidation;

3) demography and education;

4) economic structure, wages and unemployment.

We also look at 'random events' on the referendum day such as rainfall and train cancellations. We discuss each group of variables in sections 3.2-3.6. Table A1 in the appendix provides summary statistics for our variables (not standardized).

Finally, in section C.1 we also describe data used for an analysis at the level of wards within four UK cities. Wards are areas of finer geographical disaggregation, essentially city quarters, with an average population of about 7,000 (compared to roughly 170,000 residents per local authority area).

Since we are engaged in a prediction exercise and not in a structural estimation of voting behaviour, we are agnostic about whether voting results are better explained by levels of predictor variables, or by changes in those variables over a longer period. Therefore, throughout the analysis whenever available we generally use both levels and changes.[10]

[9] Due to missing covariates, we drop Northern Ireland and Gibraltar from the available maximum of 382 areas. A few covariates are also missing for some additional local authority areas, which is why some specifications in our regression tables contain fewer observations.

[10] As a robustness check, we use levels and changes separately in Tables A3 and A4 in the appendix.

3.1 Voting outcomes

We collect data on turnout and vote shares at the local authority level for the 2016 EU referendum held on 23 June 2016. Vote Leave won 51.9 percent of votes in the EU referendum, with a standard deviation of 10.4 percent across UK local authority areas. 46.5 million voters were registered in total, and 72.2 percent of these turned out. Thus, 17.4 million voted for Leave and 16.1 million for Remain. These numbers correspond to 37.4 percent and 34.7 percent of eligible voters, respectively.

Figure 1 presents a map of the support for the Leave side across local authority areas, while Figure 2 presents the map pertaining to turnout. One striking observation is that some urban centres seemed to have particularly low turnout. Within London, six local authority areas (the City of Westminster along with the Boroughs of Newham, Camden, Lewisham, Tower Hamlets, Barking and Dagenham) had turnout of less than 65 percent (out of a total of only 22 local authority areas across the whole of the UK). Since support for Remain in the EU was strongest in London, low turnout could potentially have affected the overall margin of the result. In section 4.4 we will discuss speculative scenarios to see how likely differential turnout is in explaining the result.

While our analysis is cross-sectional in nature, it is interesting to note that the 2016 EU referendum result is closely correlated with the UKIP vote share in the 2014 European Parliament elections, as illustrated in Figure A2 in the appendix.[11] The positive relationship is striking. A simple regression line has an intercept of around 25 percent and a slope close to unity, yielding an R^2 of 75 percent.[12] While it is beyond the scope of our correlational analysis to uncover the true causal relationships, the tight link suggests that the evolution of UKIP support over time may provide a lens for understanding the causal drivers behind the EU referendum result (see Becker and Fetzer, 2016 for an analysis of UKIP vote shares in EP elections in 1999, 2004, 2009 and 2014).

3.2 EU exposure: immigration, trade and EU transfers

In a referendum on EU membership, the most natural predictors for the decision to remain in or leave the EU are variables that capture the UK's exposure to the EU. Depending on the costs and benefits from EU membership that different parts of the country perceive, measures of immigration, trade and receipt of EU structural funds is likely to matter for the Vote Leave share.

[11] Also see Goodwin and Heath (2016).

[12] In the working paper version of this paper, Becker et al. (2016), we also used UKIP vote shares in the regression analysis.

Immigration We first consider immigration, a central topic throughout the Leave campaign. In the wake of the Eastern enlargement of the European Union in 2004, the UK, Ireland and Sweden were the only countries not to impose transitional controls on migrants from new member states. The UK only put in place immigration controls when Bulgaria and Romania joined the EU in 2007, but those elapsed by 2014. Given that UK wages are a multiple of those in accession countries, many Eastern European workers moved to the UK, and immigration has been at the forefront of the public debate ever since, especially in the tabloid press. While net immigration from the EU to the UK was only 15,000 in 2003, in the year before Eastern enlargement, it jumped to 87,000 in 2004. It fell slightly in the aftermath of the global financial crisis when pound sterling depreciated, only to rise strongly again to an all-time peak of 184,000 in 2015.[13] Nevertheless, it comes as a surprise to many political observers that the *net* migrant stock with other EU countries is substantially *lower* in the UK than in Germany, Spain and France, not least because the UK has a fairly high emigration rate to the EU compared to these countries (Vargas-Silva, 2012).

In fact, immigration has ranked as a top priority for UK voters over the last decade, together with the economy and the National Health Service (NHS). A key pillar of the Leave campaign was to promise control of immigration by restricting the free movement of labour from other EU countries. However, throughout that period net immigration from non-EU countries always exceeded EU net immigration typically by a substantial margin, especially prior to 2013 (see Wadsworth et al., 2016).[14]

To capture the trends in immigration, we link data from the 2001 and 2011 censuses on levels as well as growth rates in the local resident shares by three origin groups (EU 15 countries, the 12 EU accession countries that joined the EU in 2004 and 2007, and non-EU migration).[15]

Trade The 'take back control' theme of the Leave campaign also extended to the free movement of goods and services. Many voters perceived international trade not as an opportunity to sell to foreign markets but rather as unwelcome competition threatening their jobs and livelihoods. To address the role played by 'globalisation' and 'foreign

[13] Figures are from the Office for National Statistics.

[14] In a string of recent immigration-related referenda in Switzerland, the rural regions that had comparatively little immigration tended to vote most strongly against it, see here. Likewise, EU migrants are heavily concentrated in London where the Remain vote share was particularly high.

[15] The migration growth rate is defined as the change in the number of migrants between 2001 and 2011 relative to the local resident population in 2001. Our migration data are by country of birth, not by citizenship. That means first-generation immigrants from earlier migration waves (e.g., from Commonwealth nations in the 1950/60s) are captured if still alive in 2011.

competition' in the context of international trade, we match data on EU trade integration of individual UK regions to local authority areas. Specifically, we measure trade integration as the share of value added in a UK region that can be attributed to consumption and investment demand in the rest of the EU. This data is available by 37 NUTS2 regions in the UK for the year 2010. There is considerable variation across UK regions. The highest degree of trade integration can be found in East Yorkshire and Northern Lincolnshire, Cumbria, Leicestershire, Rutland and Northamptonshire (over 14 percent), and the lowest in Inner London, North Eastern Scotland, Eastern Scotland and the Highlands and Islands (around 4 percent).[16] We stress that for the purposes of interpreting our regression results in section 4, it is important to keep in mind that due to the higher aggregation at the NUTS2 level, we have in principle less variation in our trade integration measure.[17]

EU transfers Lastly, a further central topic of the referendum campaign was the size of British EU budgetary contributions. The Leave campaign quoted a figure suggesting that every week, £350 million were sent to Brussels as the UK's contribution to the EU budget. This figure was widely criticized as misleading since a significant share of the funds were returned to the UK (the net contribution was closer to £120 million per week).[18] While the gross payment towards the EU budget is not attributable to voting areas, we can track funding received from the EU. Data on EU funding is available by 133 regions in the UK. Those are essentially NUTS3 regions but were aggregated in a few cases because of past changes to boundaries of NUTS3 regions. We map them onto the local authority areas. On the one hand, EU funding has been found to be generally beneficial to regional growth (Becker et al., 2010, 2012, 2013). But on the other hand, EU funding may be perceived by voters as a handout and a symbol of foreign dependence (Davies, 2016).

3.3 Public service provision and fiscal consolidation

The referendum also presented an opportunity for those 'left behind' to express their anger, more generally speaking. The Vote Leave promise of 'taking back control' lent itself to an interpretation beyond control of borders and was seen as invitation to take

[16] We source the data on value added shares from Los et al. (2017). It combines the contributions of all major sectors to regional GDP (services, manufacturing, construction and primary industries including agriculture, mining and energy supply). Los et al. (2017) find a *positive* correlation between EU trade integration and the share of voters intending to vote Leave.

[17] See Arnorsson and Zoega (2016) for an analysis of the Brexit vote at the level of those NUTS2 regions. However, these authors do not use any trade-related covariates.

[18] The £350 million number is even incorrect as a gross figure since it does not account for the UK rebate.

back control of their own lives and express anger over a ruling class that has not addressed reduction congestion of public services, whether or not related to immigration.

Fiscal cuts In the wake of the global financial crisis, the coalition government brought in wide-ranging austerity measures to reduce government spending and the fiscal deficit. At the level of local authorities, spending per person fell by 23.4 percent in real terms from 2009/10 until 2014/15. But the extent of cuts varied dramatically across local authorities, ranging from 46.3 percent to 6.2 percent with the sharpest cuts typically in the poorest areas (Innes and Tetlow, 2015). It is important to note that the variation of cuts across local authorities is driven by the unequal share of the population that receives different kinds of benefits, hence cuts are generally larger in more deprived areas. Given this, it is not surprising that in regressions where we control for demographic characteristics that capture 'need', the fiscal cuts coefficient changes substantially, reflecting the more fundamental nature of the underlying demographics that are themselves predictors of those cuts. While some spending budgets such as the NHS were ring-fenced and therefore experienced small or no cuts, other areas such as social services and housing benefits faced drastic spending reductions. At the same time, a growing population and immigration further increased pressure on public services.

We obtain data compiled by the Financial Times capturing the geographic heterogeneity of budget cuts across all UK local authority areas. These variables capture various spending cuts affecting housing benefits, non-dependant deductions, disability living allowance, incapacity benefits, child benefits and tax credits. The measures are expressed in terms of the financial loss per working adult in pounds sterling per year over the period from 2010 to 2015. The overall financial loss per working adult varies between £914 in Blackpool and £177 in the City of London. Most fiscal cuts were applied across the board affecting individual claimants across the country fairly homogeneously. This implies that the geographic variation in the size of the fiscal cuts captures the underlying baseline degree of demand for benefits: the places with highest demand for benefits were naturally more affected.[19] In other words, fiscal cuts largely reflected (and reinforced) weak fundamentals (see also Beatty and Fothergill, 2016).

NHS service delivery The Leave campaign made frequent reference to the pressure on public services in general and the NHS in particular, mainly holding immigration responsible although in fact, immigrants from the EU were net contributors and thus

[19]The data is available here and explained in more detail here.

subsidized public spending and helped to reduce the fiscal deficit (Wadsworth et al., 2016).

As a measure of NHS service delivery we capture the fraction of suspected cancer patients who are being treated within 62 days from being first seen by a doctor. This is a key NHS health target metric for which we obtained data for the fourth quarter of 2015/16 across England, Scotland and Wales.[20] We match the local authority areas to 230 clinical commission groups under the oversight of the NHS Commissioning Board Authority. The fraction of treated patients varies from around 60 percent to 90 percent.[21]

Pressure on the housing market Immigration is often made responsible for pressures on the housing market, which is suffering from a structural deficit of newly built properties especially across the growing urban centres in the South. We therefore complement the fiscal consolidation and NHS waiting time variables with data from the 2001 and 2011 censuses on the shares of the population owning a house (outright or mortgaged), or living in council-provided rental housing.

Commuting In addition, we use 2011 census data to control for the share of working age residents that commute to Inner London for work. Commuting is supposed to capture two things: first, it can be seen as 'lack of job opportunities' at place of residence. Second, it measure the luxury enjoyed by those with well-paid jobs in London who reside in posh suburbs. The effect of this variable on the Vote Leave share is ex ante unclear.

Public sector jobs Furthermore, we consider the public employment share as measured by the Business Register and Employment Survey. This is another important measure of local service provision and jobs under threat in the light of austerity policies.

3.4 Demography, education and life satisfaction

It has been argued that older voters were more prone to Vote Leave, while younger voters overwhelmingly supported Remain. Also, less educated voters are those who

[20] The NHS publishes waiting times for a host of potential treatments, but the data for suspected cancer patients were by far the most complete and constitute a treatment that is of particular urgency where prolonged waiting times can have life-threatening consequences.

[21] We compute the average within a local authority area. If no clinical commission group sits in a local authority area, we take the value of the nearest one. Patients might choose not to receive treatment (unobservable to us), thereby affecting the overall fraction of treated patients.

might find it harder to grasp the opportunities from globalization in the form of EU membership and at the same time suffer most from the challenges posed by globalization. Voters dissatisfied with their lives and or regions with large disparities in life satisfaction may have been more prone to Vote Leave. We try to capture those factors as follows.

Age structure To reflect characteristics of the local population, we rely on data from the 2001 and 2011 censuses on the share of the local population by age brackets.[22]

Education We capture the education of the local population by the shares of people with various qualification levels.[23] Figure A3 in the appendix provides a map of the population shares with no qualifications in the year 2001. We note that, to the extent that education and the age structure of the population are more fundamental factors, it will not be surprising to find that they pick up some of the variation in other 'intermediate' predictor variables of the Vote Leave share: as argued above, fiscal cuts were largely fiscal cuts to benefits enjoyed by older and less educated parts of the population. Also, migration from Eastern Europe was largely into less educated areas (see Becker and Fetzer, 2016), so again we expect variation in education to affect the coefficients on migration variables when all of those variables are pooled in the same regression.

Life satisfaction We obtained so-called 'headline estimates' of personal well-being from the Annual Population Survey (APS) provided by the Office of National Statistics, available at the level of local authorities, for the year stretching from April 2015 to March 2016. We use both the mean life satisfaction as well as the coefficient of variation over the four categories Low, Medium, High, and Very High.

3.5 Economic structure, wages and unemployment

A typical narrative is that the Leave campaign resonated particularly well with voters in areas that had experienced prolonged economic decline, especially in the manufacturing sector. Those at the lower end of the wage distribution might have been more

[22] Those brackets are under the age of 30, between 30 and 44, between 45 and 59, 60 and older. We ultimately use the share variable for the age group 60 and older as our reference group. As discussed already, best subset selection – while powerful – is also prone to a curse of dimensionality problem so that we cannot use an endless number of covariates.

[23] There are in principle five brackets: no qualifications, level 1 (up to 4 GCSEs or equivalent), level 2 (5 or more GCSEs or equivalent), level 3 (2 or more A levels or equivalent) and level 4+ (undergraduate degree, professional qualification or equivalent). We ultimately use share variables for the lowest and highest qualification levels, the remainder being the reference group.

prone to competition from Eastern European migrants, so wages are also a potentially important predictor.

Sector structure To capture the economic structure across local authority areas we collect data on the employment shares in retail, manufacturing, construction and finance in 2001 and 2011. We use both the employment shares across those sectors in 2001 as well as the changes in those shares between 2001 and 2011 as predictor variables.

Wages We add information on wages and earnings obtained from the Annual Survey of Hours and Earnings. Specifically, we focus on levels for the year 2005 and changes in median wages between 2005 and 2015.[24] Similarly, we include data from the Annual Population Survey/Labour Force Survey, in particular the unemployment rate, the self-employment rate and overall participation rate of the working age population.

3.6 Campaigning and events on the referendum day

Apart from the four broad groups of predictor variables listed so far, events on the day of the poll may also be important in explaining turnout and voting patterns. Heavy rain in London and the Southeast of England led to the cancellation of trains during the evening rush hour, and a number of commuters did not reach the voting booths in time before their 10pm closure. In line with earlier research (see Madestam et al., 2013; Meier et al., 2016), this weather pattern may potentially influence turnout and the voting result in affected areas. We pair daily rainfall measurements from the CHIRPS precipitation data set, available at a 0.05 degree resolution, with the share of residents in a local authority area who commute to London. We investigate whether significant rainfall had an effect on turnout and the Vote Leave result across local authority areas that host a large share of London commuters.[25]

In addition, we also study the role of the tabloid press. We construct a measure covering the extent to which the Daily Mail, the Sun and the Daily Express are read by residents in these areas. For lack of detailed geographic circulation data, we rely on the British Election Study (BES) data for 2001, 2005, 2010 and 2015. All these surveys contain a question whether an individual reads a daily newspaper and if so, which one it is. We match respondents (who live in wards of sampled constituencies) to the local authority area and compute an average of the number of respondents over all these BES

[24] Bell and Machin (2016) report a negative relationship between median wages and the Vote Leave share.
[25] The CHIRPS data is available here.

surveys who report reading the Daily Mail, the Sun and the Daily Express.[26] These are naturally noisy proxies and they are only available for around 185 local authority areas, which is why we treat this analysis as a separate exercise.

4 Results

In section 3, we discussed our variables in different groups. To get a first indication of how these groups are related to the 2016 EU referendum result, in section 4.1 we first regress the vote shares separately on the variables of each group. Our aim is twofold. First, discussing groups of variables separately allows us to concentrate on the relative importance of variables within a thematic group as predictors of the Vote Leave result. In Appendix B we also perform speculative back-of-the-envelope calculations to see by how much important predictor variables would have had to be different to overturn the referendum result. Second, looking at the R^2 for groups of variables informs us about the predictive power of thematic groups relative to each other. After this, in section 4.2 we pool the groups of variables and perform the best subset selection procedure more generally. Finally, in section 4.3 we highlight the role played by the interaction of key predictor variables. This allows us to answer questions such as whether fiscal cuts affected the referendum result more in regions with weaker fundamentals.

4.1 Predicting the Brexit vote by variable group

All of the four tables pertaining to results for the four groups of predictor variables (Tables 1-4) follow the same logic: the first column shows the one variable that has the best predictive power among all variables in the variable group. The subsequent columns show the different best subsets for regressions with two regressors (column 2), three regressors (column 3), etc. The last column reports the full set of regressors.

It is important to remember that the best subset of $k-1$ predictors is not necessarily nested in the best subset of k predictors. Table 2 is a case in point where the regressor in column 1 does not appear in column 2. For this reason, in Tables 1 to 4 there is no 'triangular' structure for the columns displaying the different best subsets. Note that we standardize all right-hand side variables to mean zero and a standard deviation of one to ease comparability of coefficient estimates. The left-hand side variable is the percentage of the Leave vote, i.e., it varies between 0 and 100.

[26] We only include local authority areas with at least ten respondents across these four surveys. Restricting the set to only include local authority areas with at least 30 respondents yields very similar results.

4.1.1 Group 1: EU exposure (immigration, trade and structural funds)

In Table 1 we correlate the Vote Leave share with measures of immigration, EU trade dependence, EU subsidies (Structural Funds) and the 1975 referendum Leave share. The variation from the initial EU 15 migrant resident share in column 1 alone generates an R^2 of 29.6 percent. Adding the measure of EU trade dependence in column 2 increases the R^2 further. These two regressors together have the largest explanatory power of any two variables in this first group of predictors, jointly explaining 42.8 percent of the variation in the referendum result. The subsequent columns add only marginally to the R^2. Overall, the full set of regressors explains 48.3 percent of the variation in the Vote Leave share. Using the Akaike information criterion (AIC) as our degree-of-freedom-adjusted measure of goodness of fit, column 6 turns out to provide the best trade-off between parsimony and overall explanatory power. This column is marked by an "X" in the row "Best Subset". All subsequent tables follow the same logic.

We use migrant resident shares in levels for the year 2001 and their growth between 2001 and 2011 for three subgroups: migrants from the 12 EU accession countries that joined in 2004 and 2007, from the initial EU 15 countries, and from non-EU countries. It turns out that migrant shares in *levels* are negatively correlated with the Brexit vote as those immigrants predominantly moved to urban areas that subsequently voted for Remain in 2016. The striking observation is that in terms of migrant share *growth*, only migration from the mainly Eastern European EU accession countries *positively* correlates with the Vote Leave share. The well-established literature studying the economic implications of migration on labour market outcomes supports the notion that there are distributional consequences of low-skilled migration putting pressure on wages for low-skilled natives (see e.g., Borjas, 2003; Cortes, 2008; Borjas and Monras, 2016). Migration from Eastern Europe, predominantly of low-skilled workers, affected areas with a lower-skilled resident population.[27] As we will see below, low skills correlate with a larger Vote Leave share.

In terms of the point estimates, their interpretation is simplified by the fact that all regressors are standardized to have mean zero and a standard deviation of one. For instance, in the best subset specification displayed in column 6, a one-standard deviation higher initial EU 15 migrant share is associated with a 3.941 percentage-point lower Vote Leave share.

In Appendix B we explore, in a speculative way, what may have happened to the

[27] Becker and Fetzer (2016) estimate the causal effect of immigration from Eastern Europe on the UKIP vote share in European Parliament elections, which, as we saw above, strongly correlates with the Vote Leave share.

EU referendum vote under alternative scenarios in which migration to the UK would have been different. We find that since the vote shares do not appear very sensitive to migration, only a large reversal of the EU accession immigration experience would have swayed the vote. We stress, however, that such speculative scenarios must be taken with a large grain of salt, not least since various regressors on the right-hand side are correlated and a causal interpretation is generally not possible.

The EU trade dependence of local authority areas is also positively correlated with the Vote Leave share. The reason is that areas with a heavy concentration of manufacturing (such as the North East of England) tend to disproportionately import from and export to European Union countries, and those areas were likely to vote Leave. This finding has been highlighted in the public discussion before: those areas most dependent on trade integration with the EU were more likely to vote Leave (see Los et al., 2017). Interestingly, shortly after the referendum when Nissan threatened to stop further investment in Sunderland (one of the areas with a large Vote Leave share), pressure mounted on Westminster to do "something" to keep Nissan on board.

EU Structural Funds per capita over the EU Programming period 2007-2013 have no predictive power. Some have argued that EU subsidies in the form of EU Structural Funds would 'buy votes.' Davies (2016) argues that EU funding may be perceived by voters as a handout and a symbol of foreign dependence. As a consequence, regions receiving more money may loathe the EU more. Interestingly, Cornwall, the area receiving the largest amount of EU Structural Funds per capita, voted Leave but immediately after the referendum (on 24 June 2016) pleaded with the UK government to continue payments after EU money runs out. Our results indicate that, on balance, EU Structural Funds do not predict the Vote Leave share.

Finally, we include matched vote shares from the 1975 EU referendum as an additional regressor. There is a strong negative association between voting Leave in 2016 and 1975, suggesting different underlying attitudes and considerations across voting areas (see section 2.4).

4.1.2 Group 2: Public service provision and fiscal consolidation

In Table 2, we observe that the share of residents in a local authority area who commute to London is a strong predictor for voting Remain.[28] This might be explained by the fact that those commuting into London are relatively high-skilled who have a larger

[28]Note that people commute to London from as far as Manchester, 200 miles from London and a two-hour train ride from city centre to city centre.

tendency to vote Remain. On the other hand, house ownership is strongly correlated with the Vote Leave share. This correlation may not be surprising as house ownership is highest amongst the older section of the population. The share of the population in rented council housing, a measure of those potentially under increased pressure from migration of largely low-skilled Eastern European migrants, also has a strong positive correlation with the Vote Leave share.

Another important predictor in this group of variables is the extent of total fiscal cuts. Local authorities experiencing more fiscal cuts are more likely to vote in favour of leaving the EU. Importantly, fiscal cuts were implemented as de-facto proportionate reductions in grants across all local authorities (Innes and Tetlow, 2015). This setup implies that reliance on central government grants is a proxy variable for deprivation, with the poorest local authorities being more likely to be hit by the cuts. This makes it impossible in the cross-section (and challenging in a panel) to distinguish the effects of poor fundamentals from the effects of fiscal cuts. With this caveat on the interpretation in mind, our results suggest that local authorities experiencing more fiscal cuts were more likely to vote in favour of leaving the EU. Given the nexus between fiscal cuts and local deprivation, we think that this pattern largely reflects pre-existing deprivation. In Appendix B we provide speculative scenarios for fiscal cuts.

In a similar manner, pressure on the public health system matters. In regions where the share of suspected cancer patients waiting for treatment for less than 62 days is larger, the Vote Leave share is lower. By symmetry, where waiting times are longer, Vote Leave gains. Finally, areas with a larger share of the workforce in public employment, a measure of (a) availability of public services and (b) public jobs, the Vote Leave share is lower. In summary, results indicate that provision of public services and the severity of fiscal cuts mattered for the referendum result. Overall, variables capturing public service provision and fiscal consolidation explain slightly more than 50 percent of the variation in the Vote Leave share.

4.1.3 Group 3: Demography and education

In Table 3, we explore whether demography and education variables predict the referendum result. As predictors, we use both the baseline levels in 2001 and the growth between 2001 and 2011 of the share of the population that has no qualifications or a high qualification, respectively. The middle qualification range is the reference group. The results indicate that a larger baseline share of the population with no qualifications is associated with a larger Vote Leave share. A stronger increase in that share between

2001 and 2011 is further associated with a higher Vote Leave share. In contrast, the share of the population that has a high qualification is associated with a lower Vote Leave share. But somewhat surprisingly, faster growth of the share with a high qualification is associated with a larger Vote Leave share. We cannot exclude that this partially captures a generally faster increase in the population, which in turn might be associated with pressure on housing and public services.

In terms of age brackets, we use the share of the population aged 60 and older, which makes those younger than 60 the reference group.[29] Both a higher baseline share of older people as well as a larger increase in their share between 2001 and 2011 predict a larger Vote Leave share. This is consistent with polls in the run-up to the referendum indicating a clear age gradient in the Vote Leave share, with younger voters intending to vote Remain and older voters intending to vote Leave.[30]

We also add life satisfaction scores from the well-being questions in the Annual Population Survey. The mean score is insignificant. However, the coefficient of variation is positively related to the Vote Leave share. This finding suggests that a higher relative dispersion of well-being across voting areas, which can be interpreted as a measure of life satisfaction inequality, has positive predictive power for the Vote Leave share.

Overall, it is striking that the demography and education group of variables has the largest predictive power of any of the groups, with an R^2 of close to 80 percent and strongly significant associations in most cases between our regressors and the Vote Leave share.

4.1.4 Group 4: Economic structure, wages and unemployment

In Table 4, we concentrate on variables characterizing the sectoral structure of voting areas, both in terms of levels in the baseline year 2001 and in terms of their changes from 2001 to 2011. We single out employment in retail, manufacturing, construction and finance, and subsume all other sectors in the residual reference category. This reference category is of course quite heterogeneous, containing sectors such as agriculture, the public sector and various service sectors. This being said, a higher share of employment in the baseline year in any of the four sectors highlighted in Table 4 is associated with a larger Vote Leave share compared to the reference category.

As for the change in employment between 2001 and 2011, a stronger increase in

[29] Note that in principle, we could use more finely grained age brackets. But in the long specifications in Section 4.2, this would run into dimensionality issues for the machine learning algorithm, as explained above.

[30] In Appendix B we provide speculative scenarios for qualifications and age.

manufacturing, construction and finance employment is associated with a higher Vote Leave share. The growth of retail employment is not significantly associated with the Vote Leave share.

We also include median hourly pay as well as the interquartile pay range as a measure of inequality, again both in terms of levels and their changes (with 2005 and 2015 as the relevant years). A higher median hourly pay in the year 2005 is not significantly related to the vote. However, a stronger increase in that variable is associated with a lower Vote Leave share, consistent with the narrative that those "left behind" were more likely to vote Leave. We mostly do not find a significant relationship for the interquartile pay range, if anything a negative relationship in levels.

Finally, we add the unemployment rate, the self-employment rate and the general labour participation rate in the year prior to the referendum. A larger unemployment rate is associated with a larger Vote Leave share, but the self-employment and participation rates have no predictive power for the Vote Leave share. Overall, variables in this group explain around 69 percent of the variation in the Vote Leave share.[31]

4.1.5 Summary of analysis of four groups of predictor variables

Overall, each of Tables 1-4 yields an R^2 of at least 48 percent with a full set of regressors. The strongest explanatory power lies with demography and education variables in Table 3. Figure 3 gives a visual overview of the goodness of fit across Tables 1-4, where as a comparison the first bar represents the explanatory power of the regression underlying column 2 in Table 5 (see below).

The analysis of variables by group mainly served the purpose of considering different aspects of the referendum result in more detail and to see how well different groups of variables perform relative to each other. But of course, it makes sense to allow all groups of variables to 'compete' against each other in a single setup. This is what we turn to in section 4.2.

4.2 Best subset selection results

In Table 5 we use the best subset selection procedure for variables across *all* groups. Column 1 displays the best subset of variables when all the "best" variables from the four separate groups of regressors are combined in one joint 'horse race.' The regressors include two migration variables, EU trade dependence, the 1975 referendum vote share,

[31] In Appendix B we provide speculative scenarios for manufacturing employment and unemployment.

fiscal cuts, various qualification variables, median pay and the unemployment rate, amongst others. Overall, we obtain an R^2 of almost 88 percent with 19 variables.

Column 2 displays a full specification including all variables without performing another round of best subset selection that yields essentially the same R^2, despite the fact that the model of column 1 is a restricted version of the model in column 2. As a comparison, columns 3 through 6 re-display estimates using only the best subsets uncovered in each of the four variable groups from the previous tables. We stress that as in previous tables, Table 5 just reports conditional correlations with no causal identification.

We need to point out one caveat when it comes to the interpretation of column 2 of Table 5. While the point estimates, coefficient signs and statistical significance of variables *within variable groups* are internally consistent when we add successive regressors (using the same procedure underlying Tables 1-4), some coefficient signs and statistical significance patterns are different in the combined model of column 2 compared to columns 3 through 6. This is not surprising per se. The differences are attributable to the tight correlation between regressors *across variable groups*. For example, in column 2 the coefficient on total fiscal cuts is negative in contrast to the positive coefficient in column 4.

In particular, the demographic variables are tightly correlated with other key variables of interest. For example, the correlation between the share of individuals with no qualifications and the fiscal cuts measure is 65 percent. Similarly, the growth in the share of individuals with low qualifications may be partly driven by low-skilled migrant growth (its correlation with EU accession migrant growth is 48 percent). Hence, it is not surprising that when we remove the qualification measures from the analysis, the coefficient patterns across fiscal cuts and EU accession migration growth remain stable (see Table A2 in the appendix in contrast to Table 5).

For completeness and as a robustness check, we also perform a best subset selection exercise focusing on variables in levels (see Table A3 in the appendix) and variables in changes (see Table A4 in the appendix). In our baseline Table 5 we have two sets of regressors. First, we have a common core of variables that are in levels only. Second, we have a set of variables for which both changes as well as baseline levels are available (mostly qualification variables and employment shares). Table A3 performs best subset selection on the first set and the subset of the second set of variables that are levels only. Table A4 performs best subset selection on the first set and the subset of the second set of variables that are changes only. Given the smaller range of variables to choose from

in each table, it is not surprising that overall explanatory power in terms of R^2 is lower in principle. But it still turns out roughly the same in the case of Table A3. For the most part, certainly in Table A3, the variables show similar patterns of magnitude and significance as in Table 5.

To understand not only the predictive but rather the causal drivers of the Brexit vote, it would seem important to analyze data in panel form. We highlight that political support for the UKIP party in previous European Parliament elections, due to its strong predictive power for the Leave vote in the 2016 referendum, might be the appropriate outcome measure to better understand the causal mechanisms by which other characteristics affect the 2016 referendum result. Becker and Fetzer (2016) provide a first attempt along those lines, studying the effect of migration from Eastern Europe on UKIP vote shares over time. It seems an important future research agenda to use plausible identification strategies and possibly micro-level data on individual voters to explain voting patterns in response to changes in socio-economic fundamentals.

Finally, we also consider the voting results separately for Scotland only. As there are only 32 voting areas in Scotland, we face lower statistical power and hence a larger number of insignificant coefficients. Nevertheless, we find broadly similar regression results in terms of signs and relative magnitudes compared to those in Tables 1-5 for the entire sample. In particular, we find similar roles for higher qualification and median pay (associated with a lower Vote Leave share) and higher manufacturing employment (associated with a higher Vote Leave share).[32] Therefore, while the intercept of support for Vote Leave is clearly lower in Scotland, we do not have evidence to suggest that the coefficient patterns (i.e., slopes) for Scotland behave very differently from those for the entire sample.

Section C.2 in Appendix C documents that similar socio-economic forces also seem to be associated with the Vote Leave result when we explore within-city variation. This suggests that the underlying associations do not just mask a divide between urban and rural areas.

4.3 Interaction terms

While we have so far concentrated on a comprehensive approach to predicting the Vote Leave share, we also want to highlight whether salient factors reinforced each other. In the debate before and after the referendum, increased migration and fiscal cuts were highlighted as two salient developments over the years preceding the vote. Arguably,

[32] We do not include those results here but they are available upon request.

migration and fiscal cuts might have had a stronger influence on the Vote Leave share when hitting areas with different pre-existing conditions. In other words, we would like to see whether the interaction of local area characteristics influenced the degree to which migration and fiscal cuts influenced the Vote Leave share. Of course, we cannot carry out such an exercise for all of the variables entering our analysis so far, so we take an eclectic approach. We look at the flow (i.e., growth) of new migrants from Eastern European EU accession countries, the flow of new migrants from "old" EU member countries and the flow of new migrants from outside the EU, as well as our measure of total fiscal cuts as "flow" variables in separate regression analyses.

Each of these flow variables are interacted with one of the following "stock" variables: the share of the population with no qualifications; the sectoral share of manufacturing; the sectoral share of finance, all three measured in 2001; the median hourly pay in 2005. The results are striking and highlight some important aspects. The main effects of the stock variables that characterize "pre-existing conditions" in the first row of Table 6 are consistent across all four different "flow" variables: the share of the population with no qualifications and the share of those in manufacturing are both associated with a larger Vote Leave share, whereas the share of those in finance and a higher median hourly pay tend to be associated with a lower Vote Leave share. Turning to the main effect of the flow variables, migration from any origin region is, if anything, negatively associated with the Vote Leave share. The main effect for fiscal cuts differs across stock variables.

Most importantly, the interaction terms, which are the main focus here, show a striking pattern. A larger flow of migrants from Eastern Europe reaching a local authority area with a larger share of unqualified people or a larger share of manufacturing workers is associated with a larger Vote Leave share, whereas the opposite is true when a large flow of migrants from Eastern Europe reaches an area with a large share of those working in finance, or an area with higher median hourly pay. In other words, initial conditions matter.

The pattern is less clear for migration flows from "old" EU 15 countries and from non-EU countries. Here, point estimates on the interaction terms are generally smaller and often insignificant. This suggests that migration from Eastern Europe, which was distinct in nature by consisting of more lower-skilled migrants, had a different effect.

Interestingly, the interaction terms of fiscal cuts with the share of unqualified or manufacturing workers are insignificant. At first sight this non-finding may be seen as surprising since anecdotally, the significant welfare reforms and cuts were politically

contentious, and the Leave campaign implicitly suggested that the UK's contributions to the EU budget should be used to fund the UK's welfare system instead. Our interpretation for this non-result is as follows: most of the cuts that were implemented by David Cameron's government were not explicitly discriminatory but rather applied homogeneously across the UK. Since the demand for benefits is strongly associated with weak fundamentals such as a workforce with low qualifications, this implies that the incidence of cuts in per capita terms is strongly correlated with these weak fundamentals. In fact, the correlation between the share of the population with no qualifications and the total fiscal cuts measure is 65 percent, indicating that there is little independent variation that may be captured by an interaction effect. However, looking at the interaction between fiscal cuts and the finance share of the workforce and the median hourly pay variable, we find that larger fiscal cuts fostered a larger Vote Leave share in areas with a smaller finance sector and lower wages.

The role of media exposure We described the fact that data on media exposure is available for only less than half of local authorities. Still, many readers will be keen on understanding the role that media exposure played for the UK referendum result. Table 7 concentrates on understanding the link between education and media exposure. Arguably, less-educated voters may be more susceptible to 'negative press' in the form of anti-EU propaganda by the likes of the Daily Mail. Column 3 of Table 7 shows that turnout is neither significantly associated with the main effect of media exposure nor with the interaction term with education in column 3. However, column 6 shows that Daily Mail/Sun/Express penetration has a positive association with the Vote Leave share, both as a main effect and even more so among the least educated.[33]

4.4 Turnout as dependent variable

While our main analysis is concerned with the Vote Leave results, it is also instructive to look at turnout as an alternative outcome.[34] Table A6 in the appendix presents those results. For the sake of brevity, we just briefly highlight a few results. Columns 1 and 2 indicate that areas which experienced strong immigration growth from EU accession

[33]For completeness, we perform a best sample selection exercise including our media variable but on a smaller sample due to the missing observations. See Table A5. Media exposure shows up as significant with a positive sign. Otherwise, results are fairly similar to the baseline findings in Table 5.

[34]Given the regional nature of our analysis, we cannot say much about the motivation of individual voters to turn out. Empirical evidence using individual-level data suggests that social norms, peer pressure and monitoring play a key role in voter participation (see e.g. Gerber et al., 2008 and DellaVigna et al., 2017). For theoretical considerations on turnout and quora see Herrera and Mattozzi (2010) and Levine and Mattozzi (2017).

countries had higher turnout. Areas that had a higher support for Leave in the 1975 referendum (which tend to be areas that were more in favour of Remain in the 2016 referendum, see Table 1) had lower turnout. On balance, the results therefore suggest that turnout was lower in those areas with a higher potential in favour of Remain.

Column 4 shows that areas with more deprivation, as measured by stronger fiscal cuts, had lower turnout. Similarly, column 6 shows that areas with higher unemployment also had lower turnout. In contrast, areas with an older population and higher wages had higher turnout (see columns 5 and 6, respectively).

Rainfall Moreover, we study the extent to which bad weather across commuting zones south of London affected the EU referendum result. Rainfall led to train cancellations and may have had an influence by disproportionately reducing turnout of voters who commute into London and may have been more likely to harbour pro-EU preferences, given the strong overall support for Remain in London. The results are presented in Table 8. Results for turnout as the dependent variable are shown in columns 1 and 2, and results for the Leave share as the dependent variable are shown in columns 3 and 4. The findings suggest that the combination of rainfall and commuting into London is indeed associated with significantly lower turnout but not with significantly different vote shares. However, given our turnout scenarios for turnout below, any reasonable change in turnout behaviour across London commuting areas would not have been sufficient to overturn the referendum result.

Speculative scenarios for turnout Finally, we consider turnout scenarios (not based on regression results). According to detailed polling conducted after the referendum, turnout for the bracket of youngest voters aged 18-24 was 64 percent.[35] This compares to turnout for the same age group of less than 50 percent on average in UK general elections since 2000, and to an average turnout in the referendum across all age groups of 72.2 percent. Turnout for voters aged 25-39 was 65 percent and thus also higher than in previous general elections but by a smaller margin. On the other end of the age spectrum, voters aged 65 and above had a turnout of 90 percent. Support for Leave steadily increased with age, rising from just 27 percent for 18-24 year-olds to 60 percent for voters aged 65 and above.[36]

Could the referendum have ended up with a victory for Remain if more young

[35] See Helm (2016) for the turnout figures by age group in the referendum and Burn-Murdoch (2016a) for turnout in previous general elections.

[36] See Ashcroft (2016) for vote shares by age group.

people had turned out? We first focus on the required increase in turnout by voters aged 18-24 only. We use population shares by age from the Office for National Statistics from 2015 in combination with the above data on turnout and Leave support by age group. The age group of 18-24 year-olds makes up around 11.3 percent of the voting population. Holding fixed population shares, Vote Leave shares and the turnout of all other age groups, we calculate that turnout amongst 18-24 year-olds would have had to be approximately 120 percent instead of 64 percent to overturn the referendum result. Clearly, this would not have been feasible.

How about a broader group of Remain voters? According to Ashcroft (2016) the bracket of voters aged 35-44 still voted Remain by 52 percent, while the next bracket of voters aged 45-54 voted majority Leave. Could the referendum have gone the other way with a higher turnout among all voters up to the age of 44? We calculate the across-the-board increase in turnout in that larger age bracket which would have been required to overturn the result. This increase would have been 32 percentage points. That is, instead of the turnout of roughly 65 percent among voters up to the age of 44, a turnout of 97 percent would have been required. Of course, this is unrealistic.

Overall, we therefore conclude that higher turnout amongst the youngest section of the voting population, or even amongst all age groups that voted majority Remain, would not have overturned the referendum outcome.

4.5 Out-of-sample prediction for the 2017 French presidential election

Since our results establish correlative patterns in the data and do not allow for a causal interpretation, the question arises as to whether our results can be useful for forecasting. In particular, can our results predict other election outcomes in an out-of-sample manner?

To address this issue we consider the 2017 French presidential election. It is interesting to compare it against the UK Brexit referendum since arguably, both votes featured strong "populist" movements. In the Brexit case, the role of UKIP was fundamental in making the referendum possible in the first place. UKIP also played a key rule during the referendum campaign. In the French case, the Front National led by Marine Le Pen was a key contender.

As the dependent variable corresponding to the Leave share in the Brexit referendum, we consider the vote share for the Front National candidate Marine Le Pen, both in the first as well as in the second round of the French presidential election in April/May

2017. We examine the Le Pen vote share at the level of 95 French départements.[37] As to the right-hand side variables, we take the variables selected in the best subset in column 1 of Table 5 as the baseline specification. Our aim is to construct the corresponding French variables as closely as possible.

Due to data limitations we have to adjust some variables as follows. The French data do not allow us to distinguish between EU 15 migrants and migrants from the 12 EU Eastern accession countries. We therefore construct an EU 27 variable that captures both groups. We also construct the corresponding UK variable. Moreover, our French wage change variable is based on average wages across French départements (median hourly pay in the UK). Instead of the interquartile pay range, we use the slightly more compressed 70th/30th percentile range in France (but we keep the interquartile variable name). A further caveat is that this data is only available at the NUTS1 region level as opposed to the département level. For the employment shares, we have to rely on the French working-age population as a denominator rather than the resident population. There are minor discrepancies in terms of the years. For instance, the French house ownership data are for 2013 (2001 for the UK). The French qualifications growth data are for 1999 and 2013 (2001 and 2011 for the UK) and are only provided across three groups (as opposed to four distinct qualification groups in the UK case). The French migrant share refers to 2008 (2001 for the UK), which is the year France fully opened its borders to migrants from the 2004 EU accession countries. In addition, we drop the 1975 referendum variable since there never was such a referendum on EU membership in France. We also drop the variables on the share of residents commuting to London, the council rented share, cancer patients and fiscal cuts since we were unable to find corresponding French data. Finally, we standardize our regressors as before, but we also standardize and thus demean the dependent variables so that they are measured in directly comparable units.

As our first step, we run regressions with these updated variables for both the UK and French samples. We report the results in Table A7. Columns 1, 3 and 5 include the full set of regressors. Columns 2, 4 and 6 show the best subset selection. For the UK specifications in columns 1 and 2 we obtain an R^2 of almost 85 percent, while for the French specifications in columns 3 through 6 we obtain an R^2 around 65 percent. There is no major difference between the first and second rounds of the French presidential election. Overall, the chosen variables therefore pass the plausibility test of explaining the French data fairly well. Given that the variables are initially chosen based on the

[37] We do not have sufficient data for the overseas départements and Corsica, which leaves us with 95 départements in mainland France.

UK data underlying Table 5, it is to be expected that we achieve higher explanatory power for the UK specifications.[38]

Many coefficients are very similar (and significant) across the UK and French specifications, in particular EU trade dependence and no qualifications. But in France unemployment played a more important role. Looking at the individual coefficients, we see that some of the strongest (absolute) magnitudes are found for the qualification variables, both in the UK and the French context. They tend to be substantially larger, for instance, compared to EU trade dependence. This pattern underlines the relative importance of education.

We now turn to our main objective, which is the out-of-sample prediction. We illustrate our results in Figure 4. Panel A is based on columns 3 and 4 of Table A7. It compares the fitted values of the Le Pen vote estimated off French data against the actual French data without imposing UK coefficient values yet (we use the first-round Le Pen shares for the purpose of Figure 4). The graph on the left-hand side is based on the best model (column 4) while the right-hand side is based on the full model (column 3), but the fit is roughly the same. Panel A serves as a benchmark in the sense that if we use the French data to predict the French outcome without any coefficient constraints, we obtain an R^2 of 63 percent in a regression of the fitted values on the actual values. In Panel B we show predicted values based on the UK coefficients applied to the French data, plotted against the actual French data. This is thus a constrained version of Panel A. We therefore obtain a lower R^2 of 33 percent, which is roughly halved. That is to say, using the model that is constrained to use the coefficients estimated off the UK data, we are still able to explain roughly 50 percent of the variation that the best empirical model could achieve based on the set of covariates we have available for France. This is the main result of the out-of-sample prediction exercise. Finally, Panel C compares the fitted values from panel A against the predicted values from Panel B. The correlation here is higher than in Panel B with an R^2 of 50 percent. This tells us that the two predictions are relatively more closely related, meaning that individual observations tend to deviate from the true observations in similar ways.

Overall, we conclude that the model we estimate for the Leave shares in the UK Brexit referendum is not purely idiosyncratic. It seems that similar factors are at work for the Le Pen vote shares in France. Naturally, the French model performs best when we estimate France-specific coefficients. When we constrain the underlying coefficients to the UK values, the explanatory power is approximately halved but clearly, there are

[38] Note that since we have standardized the dependent variables in Table A7, the coefficients are not directly comparable to those in Table 5.

systematic similarities between the UK and French votes.

5 Summary and policy conclusions

Policy conclusions In terms of policy conclusions, we argue that the voting outcome was driven by long-standing fundamental determinants, most importantly those that make it harder to deal with the challenges of economic and social change. They include a population that is less educated, older and confronted with below-average public services. A complex picture arises about the challenges of adapting to social and economic change – challenges that differ across local authority areas. These spatial disparities might be reinforced by people self-selecting into local areas that better fit their outlook on life, for instance socially liberal professionals concentrating in London. This self-selection might explain the perceived increase in political polarization between 'cosmopolitan' and 'provincial' areas (Jennings and Stoker, 2016).

As economic change is often driven by global trend and developments, it is in our view an important avenue for future research to better understand the relationship and interplay between domestic and international politics, in particular in the context of the supranational institutions such as the European Union. Rodrik (2016) highlights the tension between democracy and ever more globalisation if national sovereignty is supposed to be maintained. Müller (2016) argues that a lack of genuine political choice can foment populism and the rise of authoritarian parties who claim that they alone can speak on behalf of the "real people" and their true interests.

Polls, betting markets and the Westminster bubble One key question remains. If the voting outcome seems relatively clear with hindsight, why did it come as such a surprise during the referendum night? Some Remain supporters highlighted the possibility of a Leave majority early on, for example the prominent Labour politician Andy Burnham from the Northwest of England as early as March 2016.[39] But the majority of journalists and politicians seem to have been caught off guard, including staff running the Remain campaign.[40] There is some evidence that when it comes to sensitive issues, individuals are more likely to reveal their true opinions if polls are double-blind. We therefore expect that the Brexit referendum (and also the U.S. election in November 2016) will have important implications for polling methods and survey methodology.

[39] Burnham warned of "too much Hampstead and not enough Hull." See here and here.
[40] See Peter Mandelson's account of the Remain campaign here and also here.

Similarly, throughout the campaign betting markets predicted the wrong outcome, typically showing a strong majority for Remain. As most money in total was wagered on Remain (although a large number of small bets were placed on Leave) and as betting markets balance the books, it is perhaps not surprising that betting markets did not get it right. The confidence in a Remain victory was also at odds with the polls, which suggested a much tighter race. In fact, analyzing 121 opinion polls in the run-up to the referendum, Clarke et al. (2016) suggest that "Leave was almost certainly ahead of Remain over the entire last month of the campaign – and possibly throughout 2016."

It is clear that a substantial subset of politicians and the media were genuinely surprised by the referendum result. This speaks to the polarization between metropolitan and other areas. We find it plausible that the 'Westminster bubble' may play a part in understanding the voting outcome, in combination with inaccurate polling. The under-representation of anti-EU parties in the British parliament is likely a crucial contributing factor to the lack of attention in the political process paid to struggling areas, especially in England and Wales. As a result of the first-past-the-post voting system, UKIP currently has no Member of Parliament in the House of Commons out of over 600, despite the fact that UKIP came first in the most recent European Parliament elections. UKIP representatives are therefore hardly in positions of political responsibility and thus largely escape media scrutiny. It may therefore be appropriate to consider ways of improving the diversity of views represented in British politics.

Could other countries follow the British and leave the EU? Leaving the European Union amounts to a major constitutional change for the UK. Given how much British politics has struggled with political decisions that are relatively minor in comparison, for instance the expansion of Heathrow Airport or the HS2 high-speed rail network, it is astonishing how such a far-reaching constitutional matter appears to have been decided by a referendum with no more than a simple majority and without an initial parliamentary debate on the same question (Kinsman, 2016).[41] These circumstances may be unique to the UK. France, for instance, requires constitutional revisions to be passed by both houses of parliament with subsequent approval through a referendum, or by a three-fifths parliamentary majority.

In any case, the UK has always had a more ambiguous relationship with the European Union, having been denied entry twice through French vetoes (see Appendix A). Margaret Thatcher negotiated the UK budget rebate in 1984. The UK opted out of the

[41] The UK Parliament only voted on the European Union Referendum Act in 2015 but at the time did not debate the substance matter of EU membership.

Euro and the Schengen Agreement and has looser arrangements regarding the Charter of Fundamental Rights and areas of freedom, security and justice.

Nevertheless, could Brexit be followed by Frexit for Grexit? Our analysis shows how the UK is characterized by stark differences across local areas in terms of the vote outcome and underlying factors such economic structure, education and immigration growth. Facing declining incomes and the challenge of adapting to a rapidly changing environment in terms of structural change and immigration, it may not be surprising that voters in some areas seized the opportunity to lash out at the established political order (O'Rourke, 2016). Similar trends of decline and structural change in parts of the economy can be observed in other EU countries. Indeed, analyzing the vote shares for the far-right leader Marine Le Pen in the 2017 French presidential election, we find similar driving forces at work. Whilst specific political circumstances may always be unique to each country, we do not see any a priori reasons to believe that it would be impossible for a similar backlash to happen elsewhere in Europe.

Whatever Brexit option the UK pursues, Britain's EU referendum can be seen as a protest from those feeling left behind and dissatisfied with the state of politics. Politicians in other European countries would be wise to heed the call.

References

Arnorsson, A. and G. Zoega (2016). On the Causes of Brexit. *CESifo Working Paper 6056*.

Arzheimer, K. (2009). Contextual Factors and the Extreme Right Vote in Western Europe, 1980-2002. *American Journal of Political Science 53*(2), 259–275.

Ashcroft, M. (2016). How the United Kingdom Voted on Thursday... and Why. *lordashcroft.com, June 2016*.

Autor, D., D. Dorn, G. Hanson, and K. Majlesi (2016). Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure. *NBER Working Paper 22637*.

Barone, G., A. D'Ignazio, G. de Blasio, and P. Naticchioni (2016). Mr. Rossi, Mr. Hu and Politics. The Role of Immigration in Shaping Natives' Voting Behavior. *Journal of Public Economics 136*, 1–13.

Beatty, C. and S. Fothergill (2016). Hitting the poorest places hardest. Technical report, Centre for Regional Economic and Social Research, Sheffield Hallam University.

Becker, S., P. Egger, and M. von Ehrlich (2010). Going NUTS: The Effect of EU Structural Funds on Regional Performance. *Journal of Public Economics 94*(9-10), 578–590.

Becker, S. O., P. H. Egger, and M. von Ehrlich (2012). Too Much of a Good Thing? On the Growth Effects of the EU's Regional Policy. *European Economic Review 56*(4), 648–668.

Becker, S. O., P. H. Egger, and M. von Ehrlich (2013). Absorptive Capacity and the Growth and Investment Effects of Regional Transfers: A Regression Discontinuity Design with Heterogeneous Treatment Effects. *American Economic Journal: Economic Policy 5*(4), 29–77.

Becker, S. O. and T. Fetzer (2016). Does Migration Cause Extreme Voting? *CAGE Working Paper 306* (October).

Becker, S. O., T. Fetzer, and D. Novy (2016). Who Voted for Brexit? A Comprehensive District-Level Analysis. CAGE Online Working Paper Series 305, Competitive Advantage in the Global Economy (CAGE).

Bell, B. and S. Machin (2016). Brexit and Wage Inequality. *VoxEU, August 2016*.

Borjas, G. (2003). The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *Quarterly Journal of Economics 118*(4), 1335–1374.

Borjas, G. and J. Monras (2016). The Labor Market Consequences of Refugee Supply Shocks. *NBER Working Paper 22656*.

Burgoon, B. (2012). Inequality and Anti-Globalization Backlash by Political Parties. *European Union Politics 14*(3), 408–435.

Burn-Murdoch, J. (2016a). Brexit: Everything You Wanted to Know about Turnout by Age at the EU Referendum. *Financial Times, 01 July 2016*.

Burn-Murdoch, J. (2016b). Brexit: Voter Turnout by Age. *Financial Times, 24 June 2016*.

Butler, D. and U. Kitzinger (1976). *The 1975 Referendum*. London: Palgrave Macmillan.

Casella, A. and A. Gelman (2008). A Simple Scheme to Improve the Efficiency of Referenda. *Journal of Public Economics 92*(10-11), 2240–2261.

Clarke, H., M. Goodwin, and P. Whiteley (2016). Leave Was Always in the Lead: Why the Polls Got the Referendum Result Wrong. *British Politics and Policy Blog, London School of Economics and Political Science, July 2016*.

Clarke, H., P. Whiteley, W. Borges, D. Sanders, and M. Stewart (2016). Modelling the Dynamics of Support for a Right-Wing Populist Party: The Case of UKIP. *Journal of Elections, Public Opinion and Parties 26*(2), 134–154.

Colantone, I. and P. Stanig (2016). Global Competition and Brexit. *Baffi Carefin Centre Research Paper Series No. 2016-44* (September).

Cortes, P. (2008). The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data. *Journal of Political Economy 116*(3), 381–422.

Crafts, N. (2016). The Growth Effects of EU Membership for the UK: A Review of the Evidence. *CAGE Working Paper 280*.

Curtice, J. (2016). Brexit Buyers' Remorse? There's More Evidence of Abstainers' Regret. *British Politics and Policy Blog, London School of Economics and Political Science, July 2016*.

Davies, W. (2016). Thoughts on the Sociology of Brexit. *Blog, Political Economy Research Centre, Goldsmiths, University of London*.

DellaVigna, S., J. List, U. Malmendier, and G. Rao (2017). Voting to Tell Others. *Review of Economic Studies 84*(1), 143–181.

Dippel, C., R. Gold, and S. Heblich (2015). Globalization and Its (Dis-)Content: Trade Shocks and Voting Behavior. *NBER Working Paper 21812*.

Dorling, D. (2016). Brexit: The Decision of a Divided Country. *BMJ*, 354.

Druckman, J. and A. Lupia (2000). Preference Formation. *Annual Review of Political Science 3*, 1–24.

Eichengreen, B. (2008). *The European Economy since 1945: Coordinated Capitalism and Beyond*. Princeton, New Jersey: Princeton University Press.

Evans, G. (2000). The Continued Significance of Class Voting. *Annual Review of Political Science 3*, 401–417.

Ferree, K., G. B. Powell, and E. Scheiner (2014). Context, Electoral Rules, and Party Systems. *Annual Review of Political Science 17*, 421–439.

Findlay, R. and K. O'Rourke (2007). *Power and Plenty: Trade, War, and the World Economy in the Second Millennium*. Princeton, New Jersey: Princeton University Press.

Ford, R. and M. Goodwin (2014). *Revolt on the Right: Explaining Support for the Radical Right in Britain*. Abingdon, Oxfordshire: Routledge.

Funk, P. and C. Gathmann (2015). Gender Gaps in Policy Making: Evidence from Direct Democracy in Switzerland. *Economic Policy 30*(81), 141–181.

Gerber, A., D. Green, and C. Larimer (2008). Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment. *American Political Science Review 102*(1), 33–48.

Goodwin, M. and O. Heath (2016). The 2016 Referendum, Brexit and the Left Behind: An Aggregate-Level Analysis of the Result. *The Political Quarterly 87*(3), 323–332.

Guiso, L., P. Sapienza, and L. Zingales (2016). Monnet's Error? *Economic Policy 31*(86), 247–297.

Guyon, I. and A. Elisseeff (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research 3*, 1157–1182.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics. New York, NY: Springer New York.

Helm, T. (2016). EU Referendum: Youth Turnout Almost Twice as High as First Thought. *TheGuardian.com, 10 July 2016*.

Herrera, H. and A. Mattozzi (2010). Quorum and Turnout in Referenda. *Journal of the European Economic Association 8*(4), 838–871.

Innes, D. and G. Tetlow (2015). Delivering Fiscal Squeeze by Cutting Local Government Spending. *Fiscal Studies 36*(3), 303–325.

Jennings, W. and G. Stoker (2016). The Bifurcation of Politics: Two Englands. *The Political Quarterly 87*(3), 372–382.

Kaufmann, E. (2016). It's NOT the Economy, Stupid: Brexit as a Story of Personal Values. *British Politics and Policy Blog, London School of Economics and Political Science, July 2016*.

Kinsman, J. (2016). A Brexit Post-Mortem: 17 Takeaways for a Fallen David Cameron. *OpenCanada.org, 04 July 2016*.

Levine, D. and A. Mattozzi (2017). Voter Turnout with Peer Punishment. *Working Paper, March 2017*.

Lewis-Beck, M. and M. Stegmaier (2000). Economic Determinants of Electoral Outcomes. *Annual Review of Political Science 3*, 183–219.

Los, B., P. McCann, J. Springford, and M. Thissen (2017). A Mismatch between Local Voting and the Local Economic Consequences of Brexit. *Regional Studies 51*(5), 786–799.

Madestam, A., D. Shoag, S. Veuger, and D. Yanagizawa-Drott (2013). Do Political Protests Work? Evidence from the Tea Party Movement. *Quarterly Journal of Economics 128*(4), 1633–1685.

Matsusaka, J. G. (1992). Economics of Direct Legislation. *Quarterly Journal of Economics 107*(2), 541–571.

Meier, A., L. Schmid, and A. Stutzer (2016). Rain, Emotions and Voting for the Status Quo. *IZA Discussion Paper 10350*.

Mullainathan, S. and J. Spiess (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Literature 31*(2), 87–106.

Müller, J. (2016). *What is Populism?* Philadelphia, Pennsylvania: University of Pennsylvania Press.

O'Rourke, K. (2016). Brexit: This Backlash Has Been a Long Time Coming. *VoxEU, August 2016*.

Rodrik, D. (2016). Brexit and the Globalization Trilemma. *Dani Rodrik's Weblog, 13 June 2016*.

Rothwell, J. and P. Diego-Rosell (2016). Explaining Nationalist Political Views: The Case of Donald Trump. *Gallup Working Paper, November 2016*.

Vargas-Silva, C. (2012). EU Migrants in Other EU Countries: An Analysis of Bilateral Migrant Stocks. *Briefing, The Migration Observatory at the University of Oxford*.

Wadsworth, J., S. Dhingra, G. Ottaviano, and J. Van Reenen (2016). Brexit and the Impact of Immigration on the UK. *Brexit Analysis No. 5, Centre for Economic Performance, London School of Economics and Political Science*.

Whitaker, R. and P. Lynch (2011). Explaining Support for the UK Independence Party at the 2009 European Parliament Elections. *Journal of Elections, Public Opinion and Parties 21*(3), 359–379.

Figures and Tables

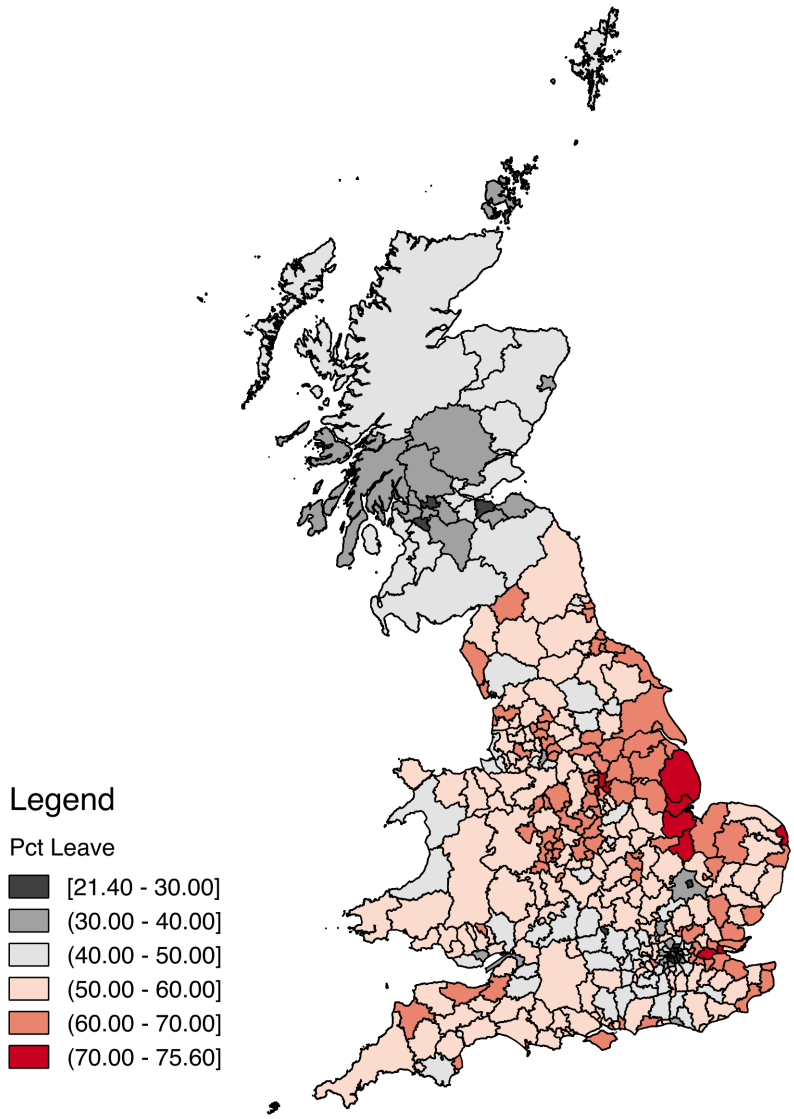


Figure 1: Map of the Leave share (in percent) across local authority areas in the 2016 EU referendum.

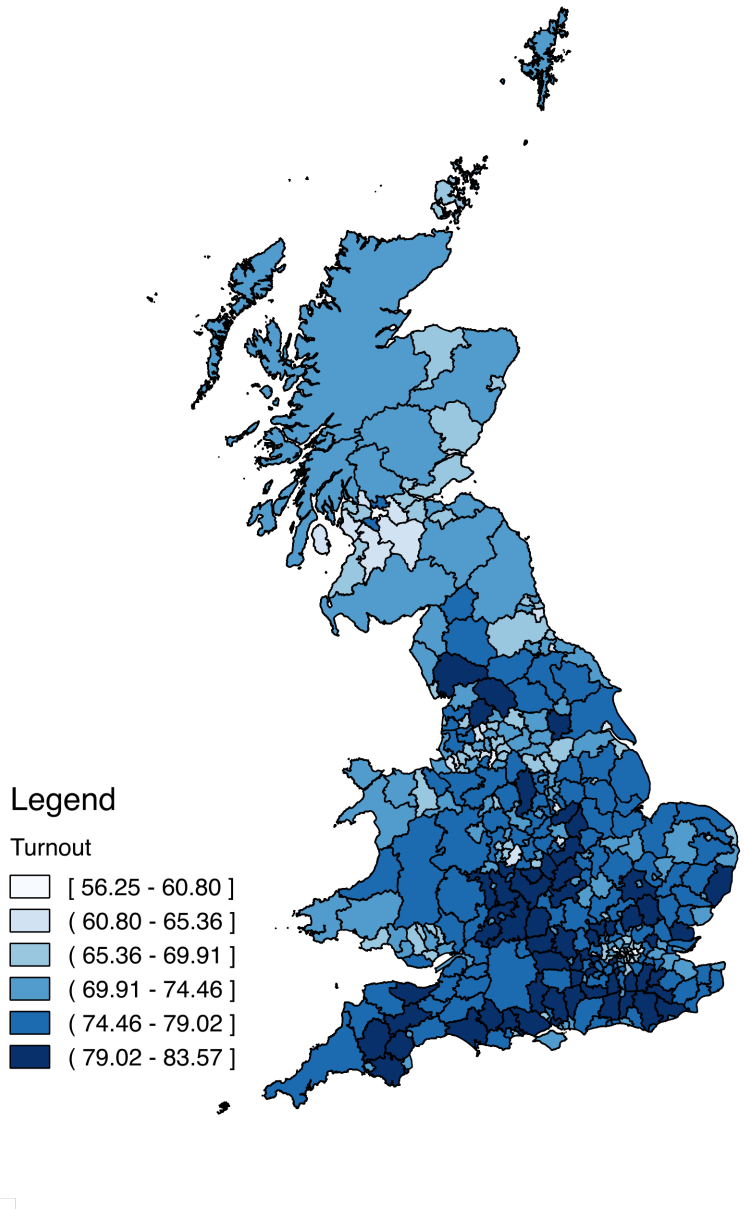


Figure 2: Map of turnout (in percent) across local authority areas in the 2016 EU referendum.

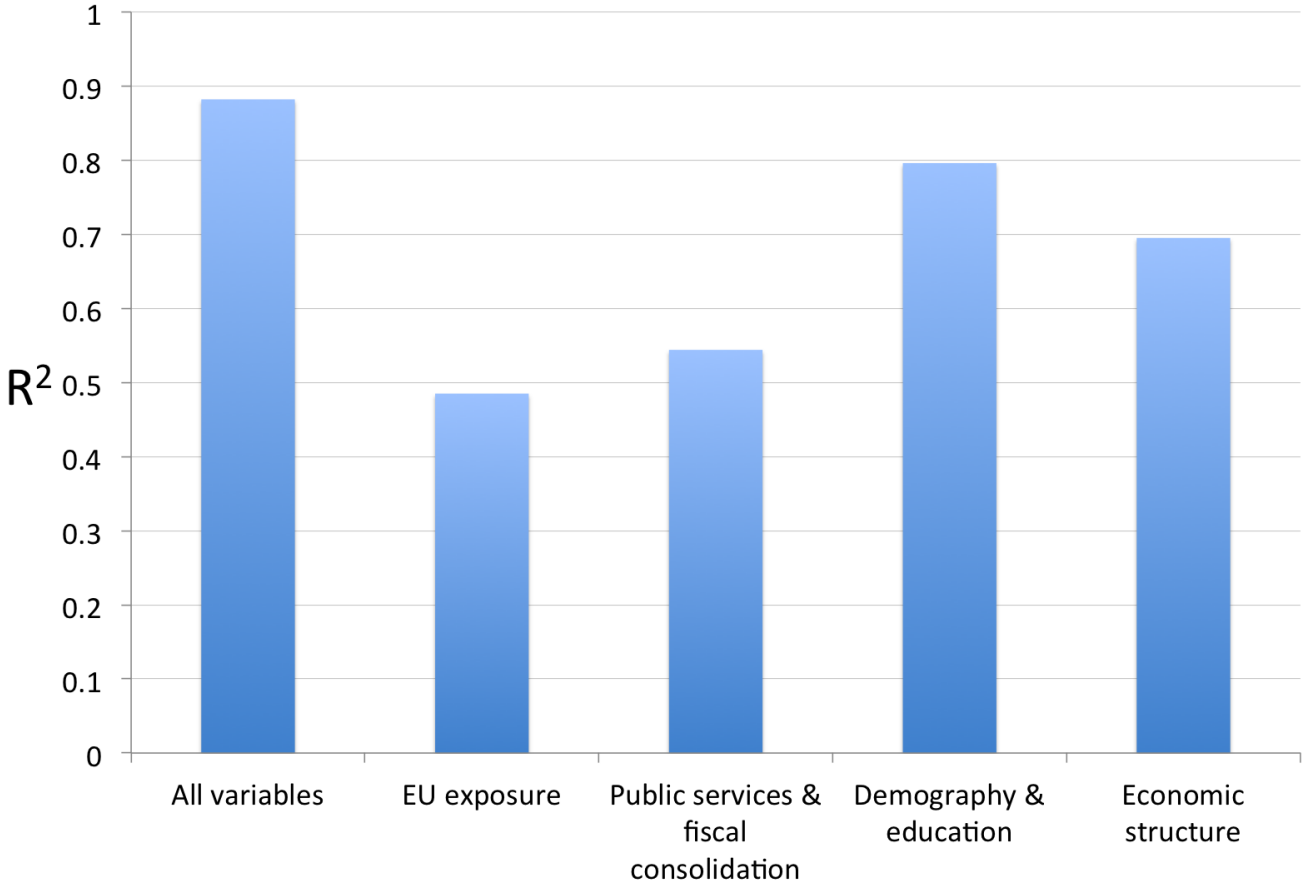
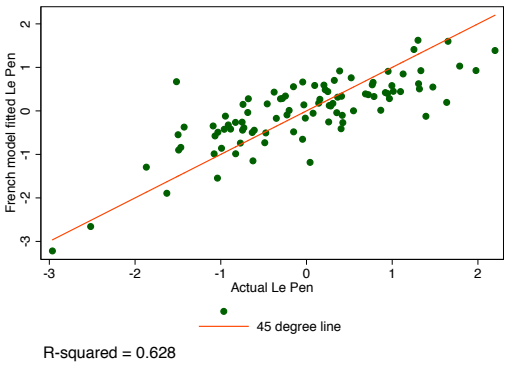


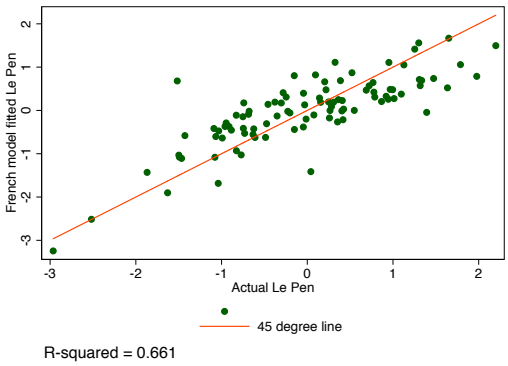
Figure 3: Goodness of fit (measured as R^2) in separate regressions explaining the Leave vote shares at the local authority area level using only regressors from the respective group of variables.

***Panel A:* Fitted values (based on French model) against actual values**

Using best French model

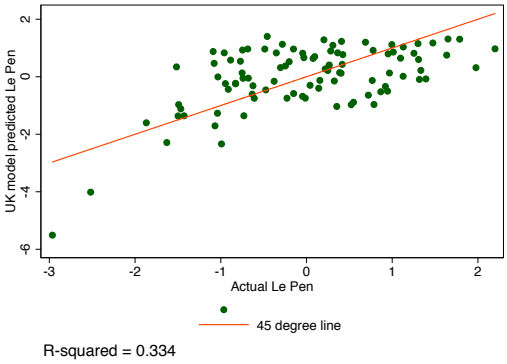


Using full French model

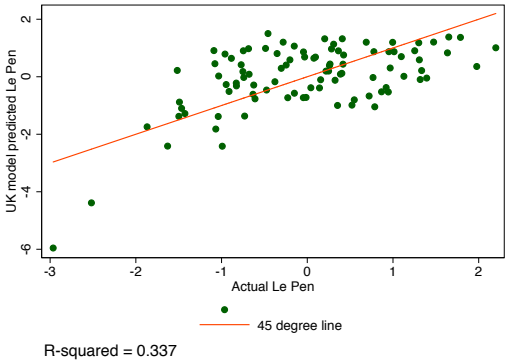


***Panel B:* Predicted values (based on UK model) against actual values**

Using best UK model

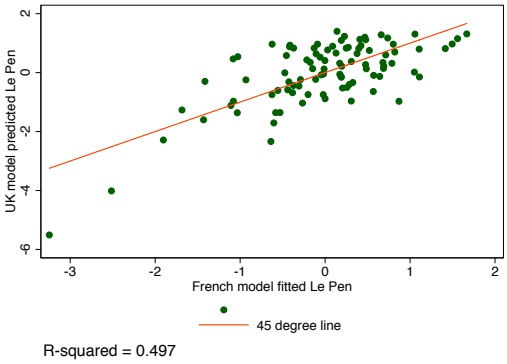


Using full UK model



***Panel C:* Predicted values against fitted values**

Using best UK model



Using full UK model

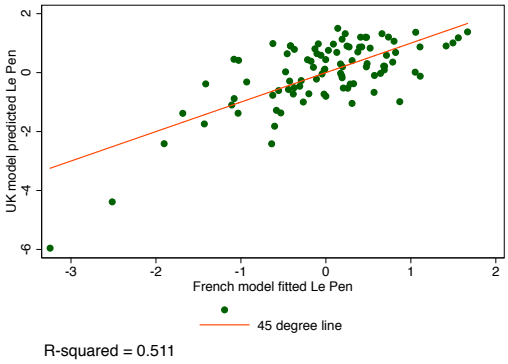


Figure 4: Results from the Le Pen vote share prediction (based on vote shares in the first round of the 2017 French presidential election). Panel A compares the fitted values of the Le Pen vote estimated off French data against the actual French data. Panel B compares the predicted values based on the UK coefficients applied to the French data against the actual French data. Panel C compares the fitted values from panel A against the predicted values from panel B.

Table 1: Predictors of Brexit Vote: EU Exposure (Immigration, Trade and Structural Funds)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Initial EU accession migrant resident share (2001)				-1.197	-1.753***	-1.651**	-1.428**	-1.267	-1.271
				(0.767)	(0.657)	(0.645)	(0.708)	(0.870)	(0.871)
EU accession migrant growth (2001-2011)					1.138**	1.376**	1.085*	1.276**	1.303*
					(0.522)	(0.533)	(0.554)	(0.632)	(0.663)
Initial EU 15 migrant resident share (2001)	-5.665***	-4.739***	-5.504***	-4.692***	-4.632***	-3.941***	-3.825***	-3.757**	-3.771***
	(0.893)	(0.854)	(1.104)	(1.361)	(1.397)	(1.518)	(1.470)	(1.475)	(1.453)
EU 15 migrant growth (2001-2011)						-1.165	-1.120	-0.921	-0.914
						(0.771)	(0.753)	(0.841)	(0.827)
Initial migrants from elsewhere resident share (2001)								-0.570	-0.504
								(0.972)	(1.223)
Migrants from elsewhere growth (2001-2011)									-0.102
									(0.859)
Total economy EU dependence (2010)		3.896***	2.586***	2.466***	2.536***	2.395***	2.659***	2.622***	2.616***
		(0.407)	(0.495)	(0.465)	(0.457)	(0.449)	(0.487)	(0.492)	(0.494)
EU Structural Funds per capita (2013)							0.556	0.525	0.522
							(0.571)	(0.575)	(0.576)
1975 referendum Leave share			-2.401***	-2.356***	-2.259***	-2.121***	-2.046***	-2.038***	-2.040***
			(0.585)	(0.586)	(0.579)	(0.592)	(0.675)	(0.678)	(0.677)
Best Subset						X			
Observations	380	380	380	380	380	380	369	369	369
R2	.296	.428	.464	.471	.48	.485	.483	.483	.483

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Best subset marked by "X". Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: Predictors of Brexit Vote: Public Service Provision and Fiscal Consolidation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of residents commuting to London (2011)	-4.767***			-2.608***	-2.990***	-2.695***	-2.708***	-2.701***
	(0.353)			(0.566)	(0.538)	(0.549)	(0.545)	(0.569)
Owned (outright + mortgage) share (2001)		7.385***	7.267***	5.378***	4.818***	6.120***	6.129***	6.128***
		(0.482)	(0.490)	(0.676)	(0.648)	(0.863)	(0.866)	(0.861)
Owned (outright + mortgage) share growth (2001-2011)								0.023
								(0.511)
Council rented share (2001)						1.609***	1.771**	1.762**
						(0.609)	(0.745)	(0.718)
Council rented share growth (2001-2011)							0.275	0.280
							(0.613)	(0.625)
Total fiscal cuts (2010-2015)		5.370***	5.556***	5.056***	5.802***	5.619***	5.629***	5.637***
		(0.450)	(0.440)	(0.466)	(0.499)	(0.488)	(0.487)	(0.501)
Share of suspected cancer patient treated within 62 Days (2015)			-2.186***	-2.654***	-2.433***	-2.398***	-2.377***	-2.381***
			(0.584)	(0.663)	(0.527)	(0.510)	(0.514)	(0.527)
Public employment share (2009)					-2.166***	-2.278***	-2.260***	-2.262***
					(0.590)	(0.583)	(0.588)	(0.579)
Best Subset						X		
Observations	376	379	378	375	375	375	375	375
R2	.215	.431	.475	.503	.535	.544	.545	.545

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Best subset marked by "X". Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Predictors of Brexit Vote: Demography and Education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of res. pop. no qualifications (2001)				4.939***	7.263***	6.467***	6.445***	6.519***
				(0.745)	(0.898)	(0.844)	(0.834)	(0.902)
Share of res. pop. no qualifications growth (2001-2011)			2.697***	4.215***	5.443***	4.900***	4.938***	4.965***
			(0.436)	(0.562)	(0.588)	(0.568)	(0.560)	(0.586)
Share of res. pop. qualification 4+ (2001)	-8.208***	-8.159***	-10.103***	-6.540***	-5.763***	-6.149***	-6.030***	-6.024***
	(0.434)	(0.399)	(0.418)	(0.785)	(0.821)	(0.703)	(0.684)	(0.688)
Share of res. pop. qualification 4+ growth (2001-2011)					2.375***	2.049***	1.956***	1.950***
					(0.465)	(0.451)	(0.455)	(0.455)
Population 60 older (2001)							0.456*	0.412
							(0.254)	(0.273)
Population 60 older growth (2001-2011)				2.815***	2.622***	2.186***	2.171***	2.117***
				(0.296)	(0.291)	(0.277)	(0.272)	(0.281)
Mean life satisfaction APS well-being data (2015)								0.135
								(0.379)
CV life satisfaction APS well-being data (2015)		2.650***	2.195***			1.369***	1.300***	1.308***
		(0.293)	(0.273)			(0.236)	(0.237)	(0.239)
Best Subset							X	
Observations	380	378	378	380	380	378	378	378
R2	.621	.687	.722	.743	.776	.795	.796	.796

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Best subset marked by "X". Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Predictors of Brexit Vote: Economic Structure, Wages and Unemployment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Retail employment share (2001)	7.019***	5.514***	4.302***	4.254***	4.097***	4.304***	4.147***	4.182***	4.045***	3.759***	3.803***	3.814***	3.821***	3.724***	3.721***
	(0.418)	(0.403)	(0.430)	(0.434)	(0.398)	(0.374)	(0.377)	(0.371)	(0.382)	(0.401)	(0.402)	(0.403)	(0.434)	(0.450)	(0.456)
Retail employment share change (2001-2011)										-0.594	-0.601	-0.593	-0.591	-0.601	-0.605
										(0.429)	(0.429)	(0.428)	(0.443)	(0.444)	(0.447)
Manufacturing employment share (2001)		3.621***	3.688***	3.516***	5.405***	5.498***	5.632***	6.051***	5.901***	5.955***	5.916***	5.901***	5.909***	5.797***	5.786***
		(0.356)	(0.302)	(0.317)	(0.509)	(0.500)	(0.498)	(0.591)	(0.625)	(0.611)	(0.613)	(0.615)	(0.650)	(0.659)	(0.687)
Manufacturing employment share change (2001-2011)					2.237***	2.478***	2.537***	2.734***	2.591***	2.317***	2.362***	2.354***	2.363***	2.319***	2.316***
					(0.546)	(0.547)	(0.540)	(0.553)	(0.537)	(0.600)	(0.603)	(0.606)	(0.661)	(0.665)	(0.665)
Construction employment share (2001)			3.220***	3.203***	3.014***	3.042***	3.304***	3.338***	3.226***	3.254***	3.328***	3.314***	3.317***	3.390***	3.391***
			(0.426)	(0.417)	(0.418)	(0.411)	(0.426)	(0.422)	(0.450)	(0.441)	(0.469)	(0.481)	(0.495)	(0.506)	(0.510)
Construction employment share change (2001-2011)				1.326***	1.529***	1.643***	1.473***	1.412***	1.414***	1.336***	1.380***	1.376***	1.375***	1.419***	1.425***
				(0.384)	(0.380)	(0.395)	(0.402)	(0.407)	(0.428)	(0.413)	(0.409)	(0.408)	(0.411)	(0.411)	(0.425)
Finance employment share (2001)								0.586	0.961**	0.945**	1.063**	1.075**	1.068**	0.988**	0.986**
								(0.429)	(0.423)	(0.419)	(0.437)	(0.440)	(0.451)	(0.459)	(0.463)
Finance employment share change (2001-2011)											0.325	0.349	0.355	0.342	0.342
											(0.430)	(0.428)	(0.443)	(0.441)	(0.441)
Median hourly pay (2005)													0.059	-0.228	-0.244
													(1.063)	(1.225)	(1.243)
Median hourly pay change (2005-2015)							-0.843**	-0.786**	-1.108***	-1.071***	-1.103***	-1.123***	-1.112***	-1.090**	-1.092**
							(0.338)	(0.331)	(0.369)	(0.371)	(0.374)	(0.366)	(0.397)	(0.481)	(0.486)
Interquartile pay range (2005)									-0.861	-0.932*	-1.094**	-1.136**	-1.175	-0.912	-0.897
									(0.535)	(0.551)	(0.551)	(0.574)	(0.866)	(1.016)	(1.038)
Interquartile pay range growth (2005-2015)														0.081	0.081
														(0.448)	(0.449)
Unemployment rate (2015)						0.827***	0.873***	0.875***	0.736**	0.692**	0.639*	0.688**	0.688**	0.707**	0.703*
						(0.303)	(0.304)	(0.302)	(0.321)	(0.325)	(0.327)	(0.344)	(0.345)	(0.352)	(0.361)
Self-employment rate (2015)															-0.027
															(0.412)
Participation rate (2015)												0.132	0.128	0.210	0.215
												(0.364)	(0.379)	(0.386)	(0.383)
Best Subset										X					
Observations	380	380	380	380	380	377	377	377	369	369	369	369	369	366	366
R2	.454	.554	.637	.653	.667	.674	.68	.682	.693	.695	.696	.696	.696	.695	.695

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Best subset marked by "X". Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Predictors of Brexit Vote: Blocked Variable Selection Approach

	Combined		Different Best Subsets			
	(1)	(2)	(3)	(4)	(5)	(6)
Initial EU accession migrant resident share (2001)	-1.678***	-1.722***	-1.651**			
	(0.530)	(0.597)	(0.645)			
EU accession migrant growth (2001-2011)		-0.501	1.376**			
		(0.425)	(0.533)			
Initial EU 15 migrant resident share (2001)	2.698***	2.820***	-3.941***			
	(0.503)	(0.554)	(1.518)			
EU 15 migrant growth (2001-2011)		-0.532	-1.165			
		(0.562)	(0.771)			
Total economy EU dependence (2010)	1.100***	0.947***	2.395***			
	(0.256)	(0.282)	(0.449)			
1975 referendum Leave share	-0.916***	-0.855**	-2.121***			
	(0.315)	(0.346)	(0.592)			
Share of residents commuting to London (2011)	0.908**	0.930*		-2.695***		
	(0.426)	(0.549)		(0.549)		
Owned (outright + mortgage) share (2001)	3.273***	2.950***		6.120***		
	(0.572)	(0.583)		(0.863)		
Council rented share (2001)	0.650*	0.608		1.609***		
	(0.381)	(0.411)		(0.609)		
Total fiscal cuts (2010-2015)	-1.463***	-1.084**		5.619***		
	(0.455)	(0.544)		(0.488)		
Share of suspected cancer patient treated within 62 Days (2015)	-0.380	-0.411		-2.398***		
	(0.282)	(0.279)		(0.510)		
Public employment share (2009)		-0.234		-2.278***		
		(0.275)		(0.583)		
Share of res. pop. no qualifications (2001)	6.024***	6.740***			6.445***	
	(0.648)	(0.904)			(0.834)	
Share of res. pop. no qualifications growth (2001-2011)	2.206***	2.715***			4.938***	
	(0.435)	(0.542)			(0.560)	
Share of res. pop. qualification 4+ (2001)	-5.897***	-4.716***			-6.030***	
	(0.793)	(1.039)			(0.684)	
Share of res. pop. qualification 4+ growth (2001-2011)		0.351			1.956***	
		(0.392)			(0.455)	
Population 60 older (2001)		-0.537			0.456*	
		(0.346)			(0.254)	
Population 60 older growth (2001-2011)		0.075			2.171***	
		(0.339)			(0.272)	
CV life satisfaction APS well-being data (2015)		0.146			1.300***	
		(0.253)			(0.237)	
Retail employment share (2001)	0.689**	0.839**				3.759***
	(0.317)	(0.391)				(0.401)
Retail employment share change (2001-2011)	-0.375	-0.177				-0.594
	(0.256)	(0.301)				(0.429)
Manufacturing employment share (2001)		0.802				5.955***
		(0.543)				(0.611)
Manufacturing employment share change (2001-2011)		0.866				2.317***
		(0.547)				(0.600)
Construction employment share (2001)		0.473				3.254***
		(0.417)				(0.441)
Construction employment share change (2001-2011)	0.664**	0.604*				1.336***
	(0.304)	(0.325)				(0.413)
Finance employment share (2001)	-0.787**	-0.573				0.945**
	(0.326)	(0.362)				(0.419)
Median hourly pay change (2005-2015)	-0.455*	-0.514**				-1.071***
	(0.235)	(0.241)				(0.371)
Interquartile pay range (2005)	0.931**	0.502				-0.932*
	(0.434)	(0.448)				(0.551)
Unemployment rate (2015)	0.475*	0.472*				0.692**
	(0.267)	(0.264)				(0.325)
Observations	366	366	380	375	378	369
R2	.879	.882	.485	.544	.796	.695

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Column 1 shows best subset across all 4 groups of variables analyzed in Tables 1 through 4. Column 2 is the full specification based on best subsets determined in Tables 1 through 4. For comparison, columns 3 through 6 re-display the optimal specifications from Tables 1 through 4. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Pairwise Interactions

	Flow: EU accession migration				Flow: EU 15 member country migration				Flow: Migration from non-EU				Flow: Total fiscal cuts			
Stock:	No Qualif.	Manufact.	Finance	Wage	No Qualif.	Manufact.	Finance	Wage	No Qualif.	Manufact.	Finance	Wage	No Qualif.	Manufact.	Finance	Wage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Stock	7.048***	5.745***	-3.223***	-4.786***	6.238***	4.832***	-1.424**	-3.415***	6.805***	5.431***	-2.284***	-4.052***	10.499***	5.909***	-3.461***	-5.687***
	(0.330)	(0.482)	(0.636)	(0.435)	(0.338)	(0.483)	(0.658)	(0.506)	(0.345)	(0.449)	(0.707)	(0.479)	(0.525)	(0.479)	(0.703)	(0.670)
Flow	-1.034***	0.229	-0.752	-0.808**	-3.924***	-1.858	-5.941***	-4.987***	-1.935***	-0.882	-2.487***	-1.791***	-4.790***	0.619	1.601***	-0.292
	(0.372)	(0.684)	(0.604)	(0.366)	(0.544)	(1.442)	(0.664)	(0.750)	(0.432)	(0.686)	(0.641)	(0.560)	(0.639)	(0.497)	(0.489)	(0.532)
Interaction	1.679***	1.411*	-2.349***	-3.048***	-0.435	1.277	0.620**	0.572**	0.187	0.960*	-0.418	-1.052**	-0.210	-0.368	-1.965**	-1.544***
	(0.277)	(0.757)	(0.661)	(0.382)	(0.312)	(0.915)	(0.272)	(0.259)	(0.329)	(0.564)	(0.502)	(0.459)	(0.373)	(0.467)	(0.759)	(0.490)
Observations	380	380	380	380	380	380	380	380	380	380	380	380	379	379	379	379
R2	.527	.346	.153	.31	.591	.433	.292	.357	.525	.348	.163	.276	.61	.326	.174	.269

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote at the local authority area level. The table presents the results for a range of interaction effects, interacting pre-determined "stock" variables measured in 2001 (for the share of households with no qualifications, the share of employment in manufacturing and finance) and in 2005 for the median wage with a range of "flow" variables capturing migration growth between 2001 to 2011 and the extent of fiscal cuts. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Tabloid Press Penetration, Education and the EU Referendum

	Turnout			Pct Leave		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily Mail/Sun/Express penetration	0.886**	0.416	0.411	5.649***	1.871***	1.745***
	(0.395)	(0.263)	(0.269)	(0.814)	(0.465)	(0.447)
Share of res. pop. no qualifications (2001)		0.410	0.379		2.981**	2.270
		(1.213)	(1.242)		(1.450)	(1.468)
Share of res. pop. qualification 1 (2001)		1.151*	1.141*		-1.694**	-1.906**
		(0.662)	(0.670)		(0.789)	(0.792)
Share of res. pop. qualification 2 (2001)		3.948***	3.949***		3.725***	3.742***
		(0.477)	(0.477)		(0.592)	(0.606)
Share of res. pop. qualification 4+ (2001)		3.379**	3.329**		-5.283***	-6.432***
		(1.543)	(1.588)		(1.891)	(1.824)
Daily Mail/Sun/Express penetration × Share of res. pop. no qualifications (2001)			0.061			1.402***
			(0.309)			(0.401)
Observations	185	185	185	185	185	185
R2	.0281	.594	.594	.255	.803	.817

Notes: Table reports results from OLS regressions. The dependent variable in columns 1 through 3 is turnout as the share of the registered electorate in a local authority area that cast their votes, while in columns 4 through 6 it is the Vote Leave share. Newspaper penetration was constructed from the British Election Study data for 2001, 2005, 2009 and 2015. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: Did Bad Weather Affect the Referendum Result?

	Turnout		Pct Leave	
	Rainfall Amount	Rainfall Top Decile	Rainfall Amount	Rainfall Top Decile
	(1)	(2)	(3)	(4)
Inner London Commuters	1.310	-0.052	-6.306***	-5.380***
	(0.834)	(0.413)	(1.266)	(0.475)
Rainfall on 23 June	1.025***	2.330**	1.584***	2.560
	(0.309)	(0.979)	(0.588)	(2.090)
Inner London Commuters x Rainfall on 23 June	-1.879***	-2.162***	0.408	0.304
	(0.455)	(0.552)	(0.718)	(0.803)
Observations	372	372	372	372
R2	.137	.07	.228	.219

Notes: Table reports results from OLS regressions. The dependent variable in columns 1 and 2 is turnout as the share of the registered electorate in a local authority area that cast its vote, while in columns 3 and 4 it is the Vote Leave share. Rainfall data is drawn from the CHIRPS rainfall data product. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix to "Who Voted for Brexit? A Comprehensive District-Level Analysis"

For Online Publication

Sascha O. Becker Thiemo Fetzer Dennis Novy

June 27, 2017

A Britain and the EU

A.1 Britain's EU history

In 1957, Belgium, France, Italy, Luxembourg, the Netherlands and West Germany signed the Treaty of Rome, which created the European Economic Community (EEC) and established a customs union. The UK negotiated access during the 1960s, but the process was interrupted twice due to French vetoes. The UK ultimately joined the EEC in 1973.

The February 1974 general election yielded a Labour minority government, which then won a majority in the October 1974 general election. Labour pledged in its February 1974 manifesto to renegotiate the terms of British accession to the EEC and then to consult the public on whether Britain should stay in the EEC on the new terms if they were acceptable to the government. A referendum on 5 June 1975 asked the electorate: "Do you think that the United Kingdom should stay in the European Community (the Common Market)?" 67.2 percent of the electorate answered 'Yes'. The 1975 referendum is described in detail by Butler and Kitzinger (1976).

The UK was instrumental in bringing about the Single Market guaranteeing the freedom of movement of goods, services, capital, and labour. Since the 1975 referendum the EEC has evolved into the central pillar of what became the European Union with the Maastricht Treaty of 1992. Further political and economic integration was formalized through the treaties of Amsterdam in 1997, Nice in 2001 and Lisbon in 2009.

On 1 May 2004, eight Eastern European countries (plus Cyprus and Malta) joined the European Union. Due to fears of migratory pressures on the social welfare systems and labour markets, many continental EU countries successfully lobbied for a phasing in of the free movement of labour. Austria and Germany, for example, imposed the maximum possible transition period, restricting the free movement of labour for seven

years from the accession date. The UK was among the few countries to allow Eastern European access to its labour market from day one.

While the UK Conservative Party campaigned for 'Remain' in the 1975 referendum, Euroscepticism grew over the years. After having negotiated restrictions to benefits for EU migrants into the UK, Prime Minister David Cameron felt compelled to hold a referendum on continued EU membership on 23 June 2016. Instead of unifying his party and rebuffing Euroscepticism as he had hoped, the vote to leave led to his resignation on 24 June 2016.

Crafts (2016) reviews the literature on the economic effects of British EU membership. He argues that UK accession to the EEC in 1973 was key to raising weak British income growth since World War II relative to other European nations. UK real income may have risen by about 8 to 10 percent due to EU accession – considerably more than had been predicted in the early 1970s by proponents of EU entry. He identifies productivity growth through increased trading opportunities, foreign direct investment and stronger competition as important dynamic mechanisms. He stresses the continuous deepening of economic integration that culminated in the Single European Act of 1986, which established the Single Market, ended capital controls and liberalized trade in services.

The Maastricht Treaty introduced further political integration. But it also paved the way for the single European currency that arguably split EU members into a core adopting the common currency and a periphery keeping their own currencies. This arrangement may be difficult to sustain over the long run, and it has led to frictions between the UK and the EU in recent years, in particular regarding the provision of financial services from outside of the Eurozone. Furthermore, the slowing pace of economic integration in the 2000s softened the growth opportunities afforded by British EU membership.

In analyzing the history of the Europe since 1945, Eichengreen (2008) highlights the pattern that political and economic integration in the EU tended to be fostered by moments of crisis, shock or deep shifts. For instance, German reunification was a trigger for further integration through the Treaty of Maastricht and institutional innovation such as the Common Foreign and Security Policy. However, the shock of the Brexit vote may be a turning point that renders political integration in the remaining EU more fragile, further driven by the rise of populist anti-EU and anti-establishment parties in many member countries.

A.2 Why did the British hold an EU referendum?

It is important to highlight that the UK's particular voting system may have contributed to the rising polarization on the issue of the UK's relationship with the European Union that culminated in the 2016 referendum. The only party that has consistently campaigned on an explicit anti-EU platform over the years is UKIP. Only founded in 1991 and taking on its current name in 1993, UKIP is a fairly new contestant on the British political scene. In the 2014 European Parliament elections it won the largest vote share, beating the Labour Party and the Conservative Party into second and third place. UKIP therefore has the ability to mobilize a large number of voters (Ford and Goodwin, 2014). But due to Britain's first-past-the-post voting system UKIP is otherwise hardly represented in national UK politics. UKIP has not a single Member of Parliament and only three representatives in the House of Lords.[1]

It can be argued that the lack of political representation in the national parliament of a large block of British voters may have contributed to the estrangement of voters from their elected representatives. It may have encouraged political entrepreneurs, in particular within the Conservative Party, to reach out to this growing political support base to improve their own position within the party, thus putting a strain on the internal cohesiveness of the Conservative Party. Many commentators have argued that internal pressure within the Conservative Party was a decisive factor in pushing David Cameron to promise an in-or-out referendum in the event the Conservative Party won the 2015 parliamentary elections, which it eventually did.

The lack of UKIP representation in the national parliament, let alone in executive positions at the national level, implied that UKIP politicians never had to deliver political outcomes at the national level and were therefore difficult to hold responsible. Yet, led by the tabloid press, the media recognized that UKIP had a popular mandate by the electorate. UKIP was therefore provided with ample coverage but relatively little scrutiny.

This mismatch of political representation and popular opinion marks a key distinguishing feature of the UK's political system: UKIP politicians have not been put to the test of navigating political compromises in order to deliver for their electorates. Since most electoral systems in Europe are based on the idea of proportional representation, the political culture on the continent tends to be more flexible to include, embrace and dilute extreme political platforms, reducing their ability to capture the political system over a single issue without ever being in a position of political responsibility and

[1] This is correct as of June 2017.

stewardship.[2]

A.3 Short-run narratives and non-economic factors

Short-run narratives Hard evidence on the potential influence of short-run phenomena on the referendum outcome seems difficult to come by. We only have suggestive evidence. For instance, as we show above, rain and train cancellations on the day of the vote are not related to the voting outcome in a way that it would have made a difference quantitatively. Second, polls in the immediate aftermath of the referendum indicated that most people would have voted in the same way even after having observed the precipitous fall of sterling. This observation casts doubt on the notion that many voters were ill-informed and therefore made a decision they afterwards regretted (commonly entertained as the 'Bregret' phenomenon). However, some regret may have been felt by abstainers (Curtice, 2016), and feelings of regret may have been growing over time.[3]

Likewise, it is hard to ascertain whether individual politicians made a key difference to the voting outcome. We have in mind the dramatic announcement in February 2016 of Boris Johnson, one of the most popular and well-known UK politicians, to support the Leave campaign. Neither is it clear whether the murder of Jo Cox, a Member of Parliament from Yorkshire, in the week before the referendum shifted the vote in any major way.

Non-economic factors We would like to stress that the fundamental factors in our paper are likely correlated with non-economic factors, for instance issues of national and class identity and social values that are best measured at the level of the individual (see Evans, 2000). Based on data from the census and the 2015 British Election Study, Kaufmann (2016) shows that a "white British working class index" is strongly correlated with the Leave vote. Favourable views of the death penalty are also correlated with the Leave vote – across all income groups. Similarly, polls by Ashcroft (2016) after the referendum indicate that Leave voters felt that life in Britain today was worse than 30 years ago, while Remain voters overwhelmingly felt the opposite.[4] Based on detailed Gallup survey data on individuals, Rothwell and Diego-Rosell (2016) make similar observations about supporters of Donald Trump's presidential campaign in the U.S. He

[2] Italy's populist 5-star movement (M5S) seems to be facing a reality check now that party representatives have been elected as mayors of cities such as Rome. One academic observer commented: "Even by Italian political standards, the M5S has made an incredible mess of things so far" (see here).

[3] As of October 2016, several months after the referendum, some polls suggest that the share of Bregret voters is on the rise, see here.

[4] Individuals were asked to agree or disagree with the following statement: "Overall, life in Britain today is worse/better than it was 30 years ago."

stresses that economic variables, household income in particular, only have mixed explanatory power. Instead, 'racial isolation' and lack of 'health and intergenerational mobility' appear as more robust predictors of Trump support.

B Speculative scenarios

B.1 A speculative scenario for migrant growth

All our regressions present multivariate correlations and have no causal interpretation. Yet, given that some variables are of key interest, we can explore, in a speculative way, what would have happened under alternative scenarios in which these variables had taken lower or higher values across the UK, holding other regressors constant. Vote Leave had 51.9 percent of votes and thus would have lost the referendum with a 1.9 percentage point lower vote share. Given our regression results we can calculate the "required" differences in these variables that would have just about led to the opposite referendum outcome. Although a causal interpretation is clearly not possible, this may give us a hint of the quantitative sensitivity of the voting result to some key variables. Instead of just considering coefficients in our regression tables, we can actually 'translate' these coefficients into more easily understandable units.

We start with the regression results on migrant share growth in column 6 of Table 1. The growth of migrants from EU accession countries is the only migrant growth significantly (and positively) related to the Vote Leave share. The corresponding coefficient stands at 1.376, meaning that – *ceteris paribus* and thus ignoring potential cross-correlations with other regressors – a one standard-deviation lower migrant growth would have been associated with a 1.376 percentage point lower Vote Leave share. The migrant growth from EU accession countries would therefore have had to be $1.381 = 1.9/1.376$ standard deviations lower to make a difference to the referendum outcome. Taken literally, this means that instead of the 1.7 percent growth as actually observed between 2001 and 2011 across local authority areas on average, the UK would have had to experience a growth of -0.6 percent, i.e., a decline in immigrants from EU accession countries.[5] Therefore, only a large reversal of the EU accession immigration experience would have swayed the vote.

Overall, such a speculative scenario must be taken with a large grain of salt, not least since various regressors on the right-hand side are correlated. We consider it as

[5]The average increase in the share of migrants from EU accession countries is 0.0170 with a standard deviation of of 0.0169. Given that the standardized growth would have had to be 1.381 standard deviations lower, the "required" growth follows as $-0.006 = 0.017 - 1.381^{*}0.0169$.

a somewhat provocative starting point for future research that should dig deeper into disentangling the causal effects of various factors on voting patters. Given the speculative nature of this exercise, there are no immediate policy implications because it is not clear that, say, an across-the-board reduction in migration from EU accession countries across all local authority areas would change the relative distribution of economic gains and losses from EU membership, let alone political sentiment.[6] Nevertheless, as we will see below, we get a sense of the relative importance of various key regressors when comparing the threshold values necessary to push the Vote Leave share under 50 percent.

B.2 A speculative scenario for fiscal cuts

We explore the key policy variable of fiscal cuts in a quantitative way. The corresponding coefficient in column 6 of Table 2 stands at 5.619, meaning that *ceteris paribus* fiscal cuts would have had to be $0.338 = 1.9/5.619$ standard deviations lower to overturn the referendum result. Given that the average fiscal cut was £448 per person, this translates into "required" cuts that would have had to be £41 lower per person.[7] We stress that the £41 result should be interpreted as a lower bound in the sense that *larger* reductions in fiscal cuts may have been required to overturn the referendum result. The reason is that areas subject to more severe fiscal cuts tend to be those with relatively more deprivation (see Innes and Tetlow, 2015). Fiscal cuts are therefore correlated with other underlying determinants of the Vote Leave share.[8]

B.3 Speculative scenarios for qualifications and age

In terms of speculative scenarios for Table 3, we focus on the share of the population with no qualifications as well as the share of the population aged 60 and older. The corresponding coefficients in column 7 are 6.445 and 0.456. The "required" predictions, *ceteris paribus* and separate for each variable, imply that the share of the population with

[6] We stress that our speculative scenario, and further speculative scenarios below, are based on a linear projection. That is, we calculate the same across-the-board "required" change in the migrant share growth across local authority areas, holding all other variables constant. Thus, our regressions and projections do not take into account potential non-linear effects. For instance, it could be that the growth of a migrant population is associated with a particularly strong Vote Leave result if the initial migrant share of the population was already elevated, or vice versa if the initial stock was low and the new migration perceived as a "shock". However, similar back-of-the-envelope calculations based on the interaction effects in Table 6 confirm the relatively weak numerical association between the Vote Leave share and immigration and thus the relatively strong "required" change.

[7] The standard deviation of fiscal cuts in the data is £122. The "required" level of cuts follows as £407 = £448 − 0.338*£122.

[8] Adding other underlying determinants related to deprivation, for instance the extent of low qualifications across the population, tends to reduce the magnitude of the fiscal cuts coefficient.

no qualifications would have had to be 33.4 percent instead of 35.4 percent, while the share of the population aged 60 and older would have had to be 3.2 percent instead of 24 percent.[9] Thus, a relatively small difference in the education pattern of the population could have led to the opposite referendum outcome, whereas the outcome would have been hardly sensitive even to large differences in senior citizen population shares. Of course, both the skill and age composition of the population are not variables that easily shift, at least not in the short and medium run.

B.4 Speculative scenarios for manufacturing employment and unemployment

We are further interested in the manufacturing employment share in the base year (with a coefficient of 5.955 in column 10 of Table 4) and the unemployment rate in 2015 (with a coefficient of 0.692).[10] *Ceteris paribus* the "required" predictions yield the following results. The manufacturing employment share would have had to be 13.4 percent instead of 15.1 percent. The unemployment rate would have had to be -0.5 percent instead of 5.3 percent. Thus, the referendum outcome may have perhaps been sensitive to the structure of employment, but not realistically with regard to unemployment.

C Within-city analysis

C.1 Within-city data

Most of the patterns we uncover in the data indicate that there is systematic variation in the intensity of the support for Vote Leave that correlates with the socio-economic fundamentals of a location. However, a first glance at the data suggests a striking divide between urban and rural areas across the UK, with support for the Remain campaign significantly stronger in urban centres, especially in London.

More generally, a potential concern with the district-level analysis is that it is subject to ecological fallacy and equates district-level voting results with individual-level voting behaviour. While we are very careful to avoid this interpretation, it is certainly helpful to look at more disaggregated data. Do the same fundamentals that we identify in the analysis across local authority areas also apply to patterns within cities? Unfortunately, city or ward level data on the EU referendum is not available across the UK. The data

[9] The standard deviation of the share of population with no qualifications stands at 6.8 percent. The standard deviation of the share of the population aged 60 and above stands at 5 percent.

[10] The corresponding standard deviations of the non-standardized variables are 5.4 percent and 2.1 percent, respectively.

has simply not been collected.[11] Yet, we managed to obtain ward level data for 107 wards across four English cities (Birmingham, Bristol, Nottingham as well as the Royal Borough of Greenwich in London). We sourced these data from local newspapers that put them together at the ward level by following the live count of votes.[12] Figure A4 shows the location of the four cities in the UK, and Figure A5 provides maps of the Vote Leave shares across their wards. Note that we cannot compute the turnout since we do not have data on the size of the electorate that is registered to vote by ward level across all cities.

We match the ward-level vote shares to cross-sectional data from the English Indices of Deprivation 2015 published by the Department for Communities and Local Government. These indices rank small neighbourhoods of approximately 1,500 inhabitants across various deprivation domains. The indices measure deprivation in income, employment, education and skills, health as well as crime. The indices rank 32,000 Lower Layer Super Output Areas (LSOAs) across England. We let a higher rank number and thus a higher regressor value indicate that a location faces more deprivation. Since wards are comprised of several LSOAs, we compute the average ward level rank in a specific domain.

C.2 Exploiting within-city variation

We perform an analysis of the EU referendum result across wards for four different UK cities for which we were able to obtain data on the referendum result (Birmingham, Bristol, Nottingham and Greenwich/London). The core purpose of this analysis is to study whether the underlying patterns that seem to drive the referendum result across local authority areas are also present when we study finer spatial variation, or whether the fact that support for the Remain side was more pronounced in many cities is driven by composition effects. The intuition is as follows: if average incomes in cities are higher, it could be that the stronger support for Remain in cities such as London masks rural versus urban differences. We will document here that when studying within-city variation of the support for Leave in the referendum, we observe very similar gradients

[11] We launched several Freedom of Information requests with various agencies, in particular the Electoral Commission, but to no avail. Our understanding is that the Act of Parliament to initiate the EU referendum specified that the counting areas at which level the results should be aggregated coincide with the administrative division of local authority areas, with no provisions made that data be published systematically at finer levels of disaggregation. It is a general challenge for work on political economy and elections in the UK that election results are only available at a very coarse spatial resolution.

[12] The data are from the Birmingham Mail, the Bristol Post, the Nottingham Post, and the 853 Blog, all of which were accessible as of 10 October 2016. Subsequently in February 2017, the BBC published data on additional wards across the UK, see here.

in proxies for fundamentals such as educational attainment. This suggests that the fact that many cities appear to have supported Remain is simply driven by a composition effect.

Unfortunately, we do not have the same breadth of explanatory variables available at the ward level. Instead, we use extremely disaggregated data on deprivation – at the level of so-called "output areas", which comprise just a few street blocks – and compute the average deprivation rank across output areas in a city ward. The main index is the composite "multiple deprivation index". It is composed of underlying deprivation indices covering the following aspects: income; employment; education and skills; health; crime.[13] We convert the data to standardize the average rank into z-scores, indicating that a higher rank means a higher degree of deprivation. Thus, when we relate the indices of deprivation to the Vote Leave share in a ward, a positive coefficient implies that worse fundamentals are associated with a higher Vote Leave share.

Since the sub-indices of the overall deprivation index are very highly correlated with each other with pairwise correlations ranging between 64 and 98 percent, we cannot separately identify effects of different sub-indices. We instead show a series of univariate regressions between the Vote Leave share at the ward level and the overall index as well as various sub-indices. We include city fixed effects throughout such that all the residual variation stems from variation across wards within cities.[14]

The results from this exercise are presented in Table A8. With the exception of crime severity, all across the board more deprivation is tightly associated with a larger Vote Leave share or, vice versa, less deprivation is tightly associated with a lower Vote Leave share. The important point to observe here again is that the strongest link with support for Leave stems from the sub-index capturing deprivation in education and skills. This is further highlighted in Figure A6. It plots the univariate correlation between the deprivation in education and skills and the support for Vote Leave after city fixed effects have been removed.

These results suggest that similar demographic and socio-economic forces were driving the support for the Leave side within as well as across cities. What may at first sight appear to be a rural-urban divide is rather a composition effect.

[13] We do not separately look at the following two sub-indices: "living environment" captures aspects such as road accidents and air quality; "barriers to housing" captures aspects such as distance to a post office, which is arguably longer in posh suburban areas but does not measure deprivation in the same way as the other sub-indices. Note, however, that both of these enter the composite "multiple deprivation index".

[14] A best subset selection exercise akin to the previous analysis across local authority areas suggests that we should include city fixed effects.

D Appendix figures and tables

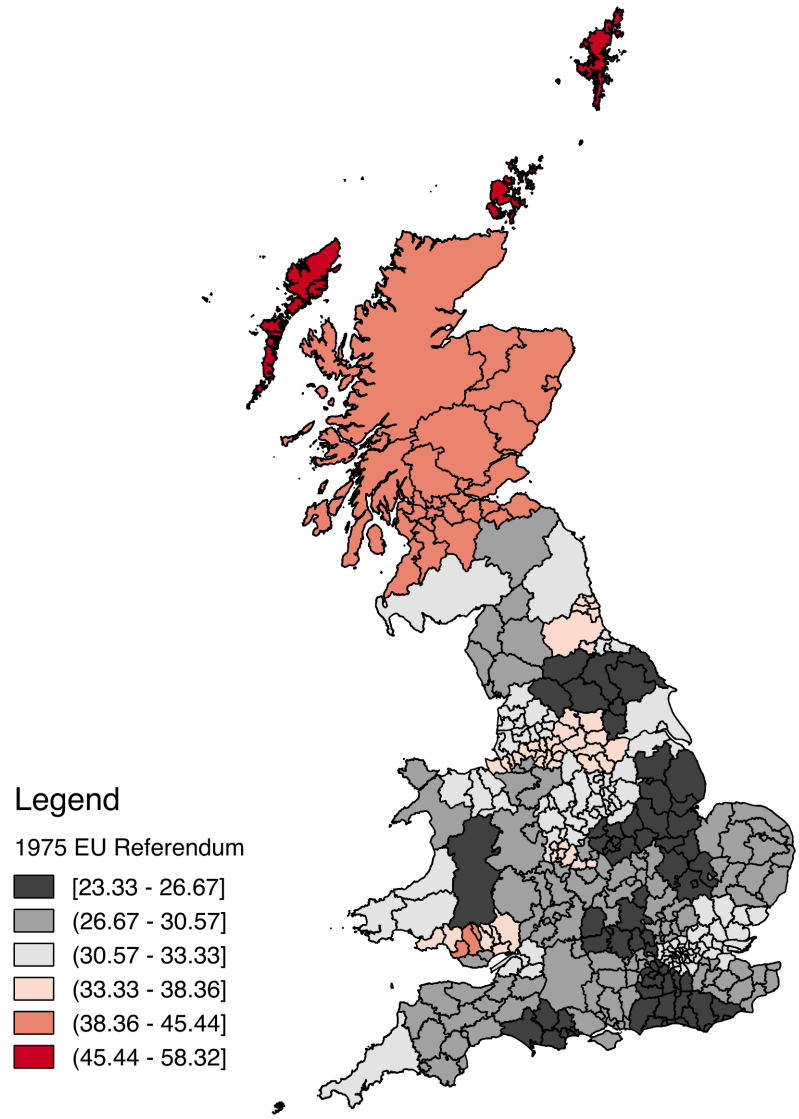


Figure A1: Map of the Leave vote in the 1975 EU referendum.

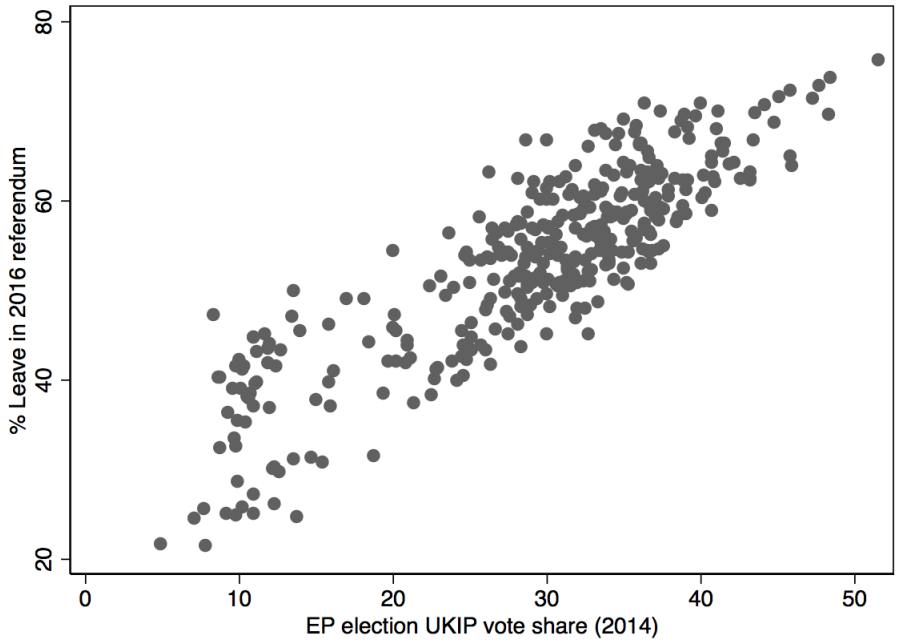


Figure A2: The figure shows the UKIP vote shares (in percent) across local authority areas in the 2014 European Parliament elections plotted against the Leave shares in the 2016 EU referendum.

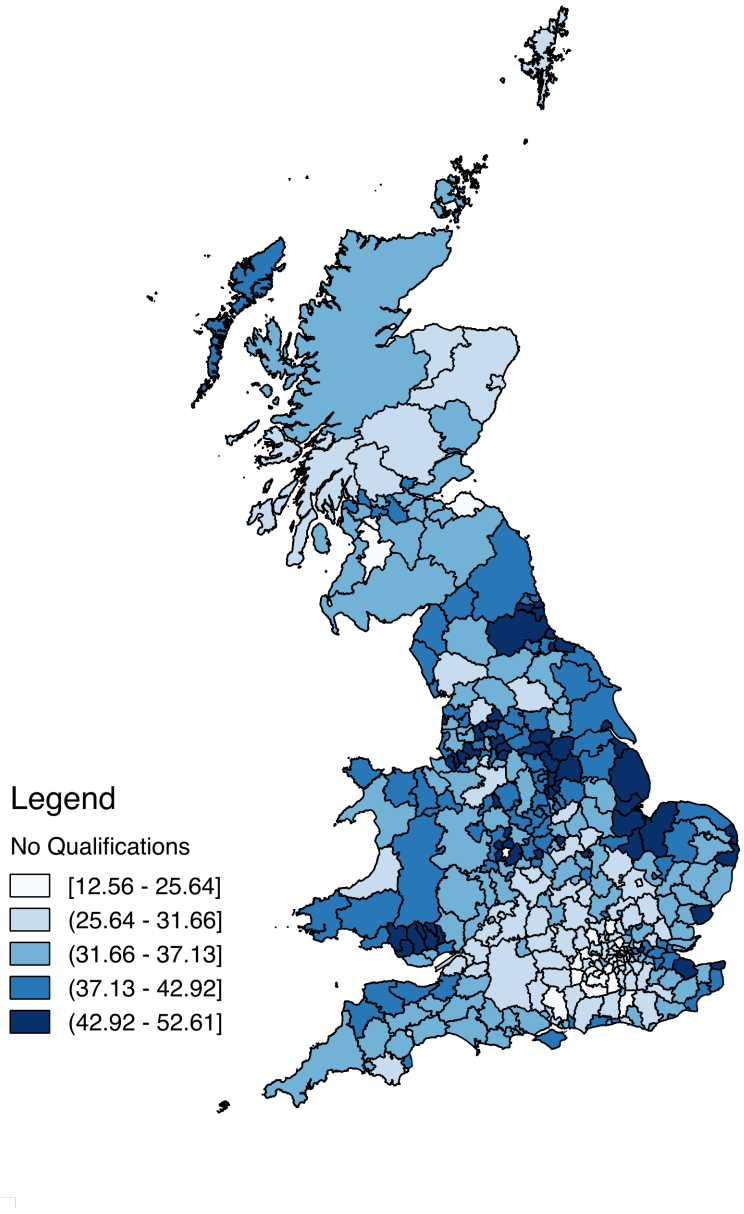


Figure A3: Map of the share of 16-74 year-olds with no qualifications across local authority areas in the year 2001.

Figure A4: Location of cities used for the ward level analysis of Leave support in the EU referendum.

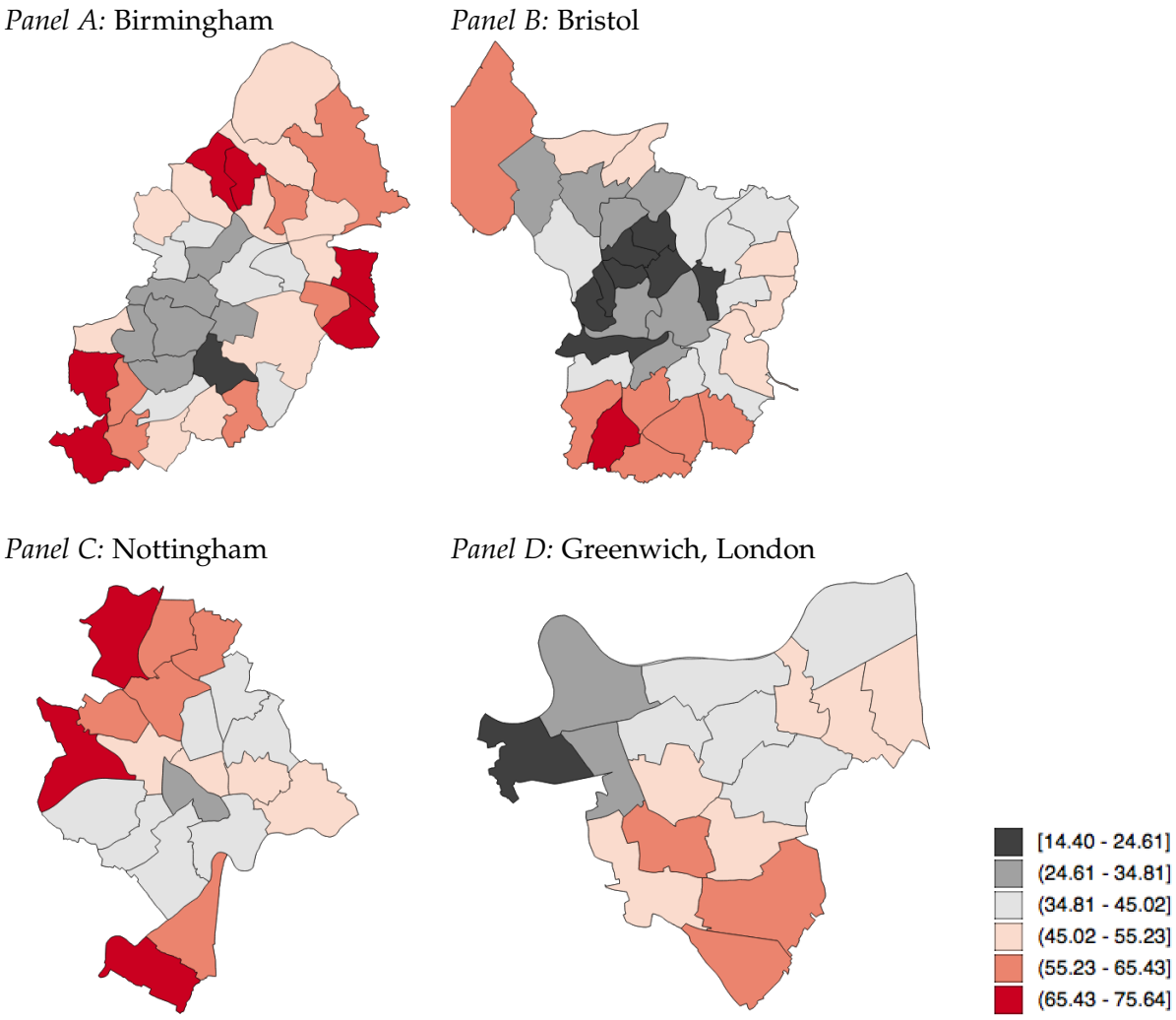


Figure A5: Maps of the Leave share (in percent) across wards in Birmingham, Bristol, Nottingham and the Royal Borough of Greenwich in London in the 2016 EU referendum.

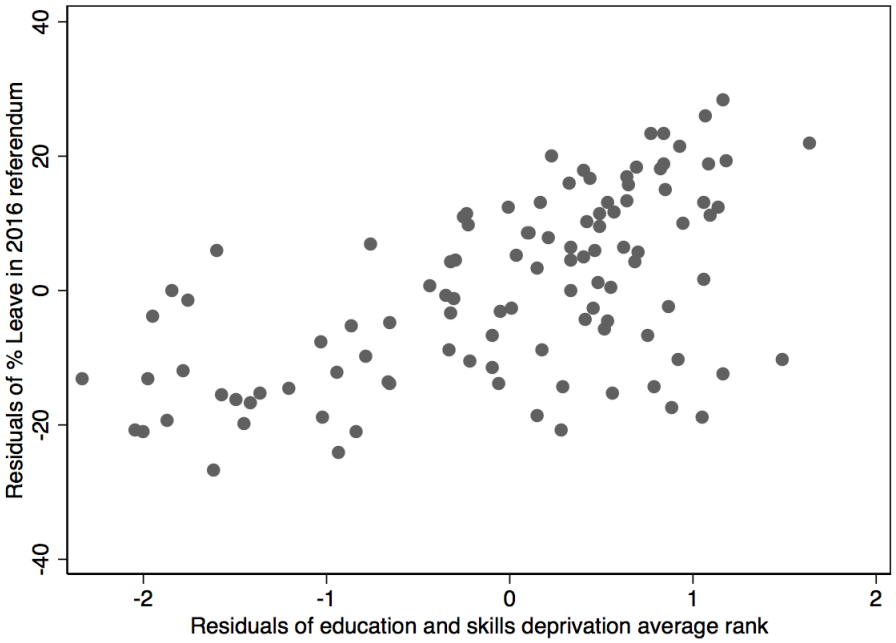


Figure A6: Relationship between the education and skills deprivation rank and the Leave share (residuals of the Leave share in percent after city fixed effects have been removed) exploiting variation across 107 wards within four cities (Birmingham, Bristol, Nottingham and Greenwich/London).

Table A1: Summary Statistics

Variable	Mean	Std. Dev.	N
Initial EU accession migrant resident share (2001)	0.002	0.002	380
EU accession migrant growth (2001-2011)	0.017	0.017	380
Initial EU 15 migrant resident share (2001)	0.013	0.011	380
EU 15 migrant growth (2001-2011)	0.003	0.007	380
Initial migrants from elsewhere resident share (2001)	0.05	0.062	380
Migrants from elsewhere growth (2001-2011)	0.025	0.03	380
Total economy EU dependence (2010)	0.097	0.032	380
EU Structural Funds per capita (2013)	46.798	96.474	369
1975 referendum Leave share	0.313	0.052	380
Share of residents commuting to London (2011)	0.017	0.044	376
Owned (outright + mortgage) share (2001)	0.698	0.096	380
Owned (outright + mortgage) share growth (2001-2011)	-0.04	0.028	380
Council rented share (2001)	0.124	0.08	380
Council rented share growth (2001-2011)	-0.038	0.039	380
Total fiscal cuts (2010-2015)	447.847	121.771	379
Share of suspected cancer patient treated within 62 Days (2015)	82.709	7.72	379
Public employment share (2009)	0.211	0.071	380
Share of res. pop. no qualifications (2001)	0.354	0.068	380
Share of res. pop. no qualifications growth (2001-2011)	-0.043	0.024	380
Share of res. pop. qualification 4+ (2001)	0.194	0.073	380
Share of res. pop. qualification 4+ growth (2001-2011)	0.075	0.015	380
Population 60 older (2001)	0.215	0.037	380
Population 60 older growth (2001-2011)	0.182	0.101	380
Mean life satisfaction APS well-being data (2015)	7.567	0.178	378
CV life satisfaction APS well-being data (2015)	1.173	0.46	378
Retail employment share (2001)	0.167	0.022	380
Retail employment share change (2001-2011)	-0.007	0.009	380
Manufacturing employment share (2001)	0.151	0.054	380
Manufacturing employment share change (2001-2011)	-0.058	0.021	380
Construction employment share (2001)	0.07	0.014	380
Construction employment share change (2001-2011)	0.01	0.006	380
Finance employment share (2001)	0.044	0.026	380
Finance employment share change (2001-2011)	-0.004	0.007	380
Median hourly pay (2005)	10.965	2.002	380
Median hourly pay change (2005-2015)	0.235	0.088	380
Interquartile pay range (2005)	9.933	3.031	371
Interquartile pay range growth (2005-2015)	0.199	0.134	367
Unemployment rate (2015)	5.26	2.108	377
Self-employment rate (2015)	10.605	3.685	378
Participation rate (2015)	78.716	4.536	379

Table A2: Removing the Qualification Variables: Variable Groups and Coefficient Signs

	Combined		Different Best Subsets			
	(1)	(2)	(3)	(4)	(5)	(6)
Initial EU accession migrant resident share (2001)	-2.305***	-2.298***	-1.651**			
	(0.364)	(0.345)	(0.645)			
EU accession migrant growth (2001-2011)		0.135	1.376**			
		(0.476)	(0.533)			
Initial EU 15 migrant resident share (2001)	1.801***	1.902***	-3.941***			
	(0.573)	(0.617)	(1.518)			
EU 15 migrant growth (2001-2011)		0.201	-1.165			
		(0.762)	(0.771)			
Total economy EU dependence (2010)	0.588*	0.677*	2.395***			
	(0.331)	(0.378)	(0.449)			
1975 referendum Leave share	-1.804***	-1.718***	-2.121***			
	(0.381)	(0.403)	(0.592)			
Share of residents commuting to London (2011)	0.648	0.413		-2.695***		
	(0.526)	(0.716)		(0.549)		
Owned (outright + mortgage) share (2001)	2.018***	2.388***		6.120***		
	(0.508)	(0.787)		(0.863)		
Council rented share (2001)		0.484		1.609***		
		(0.574)		(0.609)		
Total fiscal cuts (2010-2015)	2.436***	2.455***		5.619***		
	(0.445)	(0.579)		(0.488)		
Share of suspected cancer patient treated within 62 Days (2015)	-1.130***	-1.097***		-2.398***		
	(0.292)	(0.284)		(0.510)		
Public employment share (2009)		-0.367		-2.278***		
		(0.341)		(0.583)		
Population 60 older (2001)		0.293			3.879***	
		(0.410)			(0.532)	
Population 60 older growth (2001-2011)		-0.080			3.574***	
		(0.483)			(0.525)	
Mean life satisfaction APS well-being data (2015)		-0.060			-4.206***	
		(0.389)			(0.493)	
CV life satisfaction APS well-being data (2015)	0.580*	0.501			2.200***	
	(0.343)	(0.352)			(0.441)	
Retail employment share (2001)	2.462***	2.257***				3.759***
	(0.366)	(0.501)				(0.401)
Retail employment share change (2001-2011)		-0.044				-0.594
		(0.361)				(0.429)
Manufacturing employment share (2001)	4.456***	4.470***				5.955***
	(0.516)	(0.634)				(0.611)
Manufacturing employment share change (2001-2011)	1.721***	1.725***				2.317***
	(0.434)	(0.596)				(0.600)
Construction employment share (2001)	3.083***	3.094***				3.254***
	(0.374)	(0.389)				(0.441)
Construction employment share change (2001-2011)	1.288***	1.344***				1.336***
	(0.365)	(0.368)				(0.413)
Finance employment share (2001)		0.199				0.945**
		(0.401)				(0.419)
Median hourly pay change (2005-2015)	-0.972***	-0.955***				-1.071***
	(0.291)	(0.300)				(0.371)
Interquartile pay range (2005)	-1.116**	-1.198**				-0.932*
	(0.500)	(0.555)				(0.551)
Unemployment rate (2015)	1.004***	0.999***				0.692**
	(0.312)	(0.326)				(0.325)
Observations	366	366	380	375	378	369
R2	.801	.803	.485	.544	.316	.695

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3: Levels Only Variables: Variable Groups and Coefficient Signs

	Combined		Different Best Subsets			
	(1)	(2)	(3)	(4)	(5)	(6)
Initial EU accession migrant resident share (2001)	-0.723	-0.805	-1.197			
	(0.824)	(0.827)	(0.767)			
Initial EU 15 migrant resident share (2001)	3.679***	3.749***	-4.692***			
	(0.510)	(0.507)	(1.361)			
Total economy EU dependence (2010)	1.048***	0.990***	2.466***			
	(0.244)	(0.272)	(0.465)			
1975 referendum Leave share	-0.953***	-1.007***	-2.356***			
	(0.324)	(0.362)	(0.586)			
Share of residents commuting to London (2011)	1.786***	1.683***		-2.695***		
	(0.491)	(0.498)		(0.549)		
Owned (outright + mortgage) share (2001)	3.974***	4.342***		6.120***		
	(0.391)	(0.504)		(0.863)		
Council rented share (2001)		0.449		1.609***		
		(0.401)		(0.609)		
Total fiscal cuts (2010-2015)	-1.082**	-0.849		5.619***		
	(0.482)	(0.548)		(0.488)		
Share of suspected cancer patient treated within 62 Days (2015)		-0.262		-2.398***		
		(0.288)		(0.510)		
Public employment share (2009)		-0.268		-2.278***		
		(0.304)		(0.583)		
Share of res. pop. no qualifications (2001)	5.246***	5.032***			1.952***	
	(0.620)	(0.674)			(0.751)	
Share of res. pop. qualification 4+ (2001)	-5.748***	-5.528***			-6.512***	
	(0.757)	(0.888)			(0.867)	
Mean life satisfaction APS well-being data (2015)		-0.151			0.516	
		(0.307)			(0.377)	
CV life satisfaction APS well-being data (2015)	0.364	0.320			2.737***	
	(0.276)	(0.275)			(0.297)	
Retail employment share (2001)	1.250***	1.073***				4.364***
	(0.341)	(0.356)				(0.418)
Manufacturing employment share (2001)		0.143				3.936***
		(0.356)				(0.365)
Construction employment share (2001)		0.160				3.302***
		(0.424)				(0.420)
Finance employment share (2001)	-0.537	-0.650*				0.612
	(0.349)	(0.354)				(0.471)
Observations	376	375	380	375	378	380
R2	.856	.858	.471	.544	.694	.639

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4: Changes Only Variables: Variable Groups and Coefficient Signs

	Combined		Different Best Subsets			
	(1)	(2)	(3)	(4)	(5)	(6)
EU 15 migrant growth (2001-2011)	-2.027**	-2.181**	-4.090***			
	(1.021)	(1.048)	(0.547)			
Total economy EU dependence (2010)	1.438***	1.653***	3.360***			
	(0.406)	(0.466)	(0.509)			
EU Structural Funds per capita (2013)		0.258	0.891			
		(0.506)	(0.600)			
1975 referendum Leave share	-2.477***	-2.537***	-0.986			
	(0.496)	(0.515)	(0.605)			
Share of residents commuting to London (2011)	-1.215	-0.867		-5.429***		
	(0.907)	(0.917)		(0.430)		
Council rented share growth (2001-2011)		0.399		0.925		
		(0.420)		(0.578)		
Total fiscal cuts (2010-2015)	3.669***	3.677***		4.479***		
	(0.618)	(0.670)		(0.497)		
Share of suspected cancer patient treated within 62 Days (2015)	-1.429***	-1.330***		-2.790***		
	(0.356)	(0.377)		(0.580)		
Public employment share (2009)	-0.952**	-0.852**		-2.804***		
	(0.399)	(0.423)		(0.640)		
Share of res. pop. no qualifications growth (2001-2011)	-2.431***	-2.445***			-3.913***	
	(0.656)	(0.714)			(0.510)	
Share of res. pop. qualification 4+ growth (2001-2011)	-1.617***	-1.531**			-1.589***	
	(0.587)	(0.606)			(0.590)	
Population 60 older growth (2001-2011)	1.703***	1.532***			2.676***	
	(0.437)	(0.470)			(0.543)	
Mean life satisfaction APS well-being data (2015)		0.110			-1.958***	
		(0.472)			(0.542)	
CV life satisfaction APS well-being data (2015)	1.673***	1.528***			3.660***	
	(0.401)	(0.433)			(0.386)	
Retail employment share change (2001-2011)	-1.688***	-1.477***				-2.454***
	(0.413)	(0.413)				(0.607)
Manufacturing employment share change (2001-2011)	-0.782	-1.017*				-4.335***
	(0.504)	(0.522)				(0.547)
Construction employment share change (2001-2011)	1.462***	1.519***				1.122
	(0.461)	(0.469)				(0.699)
Finance employment share change (2001-2011)	-0.953*	-1.001*				-2.639***
	(0.531)	(0.538)				(0.504)
Self-employment rate (2015)	-1.000**	-0.942**				-1.050*
	(0.391)	(0.398)				(0.581)
Observations	375	366	369	375	378	378
R2	.68	.674	.39	.44	.343	.258

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A5: Including the Tabloid Press: Variable Groups and Coefficient Signs

	Combined		Different Best Subsets			
	(1)	(2)	(3)	(4)	(5)	(6)
Initial EU accession migrant resident share (2001)	-2.060***	-2.181***	-1.397**			
	(0.429)	(0.398)	(0.688)			
EU accession migrant growth (2001-2011)		-0.235	1.148			
		(0.632)	(0.777)			
Initial EU 15 migrant resident share (2001)	1.553*	1.509	-4.781**			
	(0.939)	(1.052)	(2.024)			
Total economy EU dependence (2010)	0.778**	0.795	2.761***			
	(0.334)	(0.487)	(0.747)			
EU Structural Funds per capita (2013)		-0.286	1.000			
		(0.443)	(0.774)			
1975 referendum Leave share	-1.475***	-1.612***	-2.175**			
	(0.510)	(0.589)	(0.940)			
Daily Mail/Sun/Express penetration	1.154***	1.133***	3.179***			
	(0.363)	(0.373)	(0.649)			
Share of residents commuting to London (2011)	1.295**	0.958		-3.077***		
	(0.498)	(0.663)		(0.725)		
Owned (outright + mortgage) share (2001)	2.879***	2.162***		5.064***		
	(0.570)	(0.812)		(0.862)		
Total fiscal cuts (2010-2015)	-1.886***	-1.278		5.936***		
	(0.684)	(0.881)		(0.685)		
Share of suspected cancer patient treated within 62 Days (2015)		-0.324		-3.577***		
		(0.394)		(0.522)		
Public employment share (2009)		0.187		-1.728**		
		(0.389)		(0.729)		
Share of res. pop. no qualifications (2001)	6.006***	5.774***			7.596***	
	(0.993)	(1.277)			(1.062)	
Share of res. pop. no qualifications growth (2001-2011)	2.448***	2.652**			5.812***	
	(0.794)	(1.179)			(0.703)	
Share of res. pop. qualification 4+ (2001)	-4.264***	-3.526***			-6.087***	
	(0.907)	(1.326)			(0.905)	
Share of res. pop. qualification 4+ growth (2001-2011)		0.211			2.538***	
		(0.526)			(0.581)	
Population 60 older (2001)		0.243			0.814**	
		(0.454)			(0.367)	
Population 60 older growth (2001-2011)		0.261			1.915***	
		(0.478)			(0.450)	
CV life satisfaction APS well-being data (2015)		-0.270			1.427***	
		(0.346)			(0.283)	
Retail employment share (2001)	1.346***	1.343**				4.321***
	(0.417)	(0.560)				(0.468)
Retail employment share change (2001-2011)		-0.360				-1.504***
		(0.402)				(0.491)
Manufacturing employment share (2001)	0.762*	1.160*				4.209***
	(0.459)	(0.667)				(0.485)
Construction employment share (2001)		0.592				3.228***
		(0.686)				(0.651)
Construction employment share change (2001-2011)	0.633	0.920**				1.493***
	(0.387)	(0.445)				(0.540)
Median hourly pay change (2005-2015)		-0.491				-0.921*
		(0.330)				(0.544)
Interquartile pay range growth (2005-2015)		0.200				0.792
		(0.367)				(0.535)
Unemployment rate (2015)	1.151***	1.225***				0.876**
	(0.412)	(0.437)				(0.430)
Observations	183	176	180	182	185	183
R2	.893	.902	.576	.57	.8	.718

Notes: Table reports results from OLS regressions. The dependent variable is the share of the Leave vote in a local authority area in England, Scotland and Wales. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A6: Predictors of Referendum Turnout: Blocked Variable Selection Approach

	Combined		Different Best Subsets			
	(1)	(2)	(3)	(4)	(5)	(6)
Initial EU accession migrant resident share (2001)	0.492***	0.469***	0.400*			
	(0.133)	(0.119)	(0.223)			
EU accession migrant growth (2001-2011)	0.517***	0.335*	-0.457**			
	(0.190)	(0.196)	(0.215)			
Initial EU 15 migrant resident share (2001)	-0.054	0.158	1.132*			
	(0.388)	(0.269)	(0.617)			
EU 15 migrant growth (2001-2011)		-0.117	-1.223***			
		(0.274)	(0.317)			
Migrants from elsewhere growth (2001-2011)	-1.037***	-1.278***	-1.733***			
	(0.236)	(0.222)	(0.281)			
1975 referendum Leave share	-0.904***	-0.709***	-2.458***			
	(0.159)	(0.164)	(0.302)			
Share of residents commuting to London (2011)	-0.349	0.045		0.406**		
	(0.249)	(0.300)		(0.202)		
Owned (outright + mortgage) share (2001)	0.865**	0.407		1.455***		
	(0.354)	(0.367)		(0.288)		
Owned (outright + mortgage) share growth (2001-2011)	0.156	-0.256		0.625***		
	(0.221)	(0.225)		(0.180)		
Council rented share (2001)		0.008		-1.248***		
		(0.199)		(0.230)		
Total fiscal cuts (2010-2015)		-1.460***		-2.341***		
		(0.274)		(0.178)		
Public employment share (2009)		-0.146		-0.782***		
		(0.152)		(0.178)		
Share of res. pop. no qualifications growth (2001-2011)		-0.335			1.145***	
		(0.230)			(0.235)	
Share of res. pop. qualification 4+ (2001)	1.091***	0.611			0.750***	
	(0.381)	(0.448)			(0.266)	
Share of res. pop. qualification 4+ growth (2001-2011)	1.474***	0.949***			2.102***	
	(0.195)	(0.174)			(0.196)	
Population 60 older (2001)	0.981***	1.183***			1.464***	
	(0.194)	(0.174)			(0.146)	
Population 60 older growth (2001-2011)	1.395***	1.004***			2.291***	
	(0.193)	(0.183)			(0.171)	
CV life satisfaction APS well-being data (2015)	0.560***	0.259**			0.847***	
	(0.125)	(0.131)			(0.124)	
Retail employment share (2001)	0.311*	0.411**				0.553**
	(0.161)	(0.163)				(0.227)
Manufacturing employment share (2001)	-0.213	-0.085				2.043***
	(0.272)	(0.258)				(0.411)
Manufacturing employment share change (2001-2011)	-0.159	-0.246				1.254***
	(0.217)	(0.220)				(0.345)
Construction employment share (2001)		-0.138				0.874***
		(0.185)				(0.207)
Finance employment share (2001)	0.181	-0.019				-0.927***
	(0.192)	(0.194)				(0.268)
Finance employment share change (2001-2011)		-0.333**				-0.953***
		(0.168)				(0.279)
Median hourly pay (2005)	0.944***	1.273***				1.561***
	(0.305)	(0.327)				(0.558)
Interquartile pay range (2005)		-0.462				1.365**
		(0.324)				(0.643)
Unemployment rate (2015)	-0.259**	0.013				-1.176***
	(0.130)	(0.113)				(0.213)
Self-employment rate (2015)		0.256*				1.466***
		(0.133)				(0.206)
Participation rate (2015)		0.274*				1.166***
		(0.145)				(0.210)
Observations	375	367	380	376	378	369
R2	.858	.886	.486	.77	.783	.621

Notes: Table reports results from OLS regressions. The dependent variable is turnout as the share of the registered electorate in a local authority area that cast its vote. Empirical models selected using best subset selection on the set of predictors using the AIC information criterion. Columns 3 through 6 display the specifications drawn from the 4 groups presented in Table 5. Column 1 shows the best subset selection across the latter 4 groups of variables. Column 2 is the full specification. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A7: UK 2016 EU Referendum versus French 2017 Presidential Election

	UK		France			
	(1)	(2)	(3)	(4)	(5)	(6)
			Share Le Pen r1	Share Le Pen r1	Share Le Pen r2	Share Le Pen r2
EU 27 migrant resident share	0.207***	0.173**	0.068		0.113	
	(0.067)	(0.068)	(0.082)		(0.078)	
Total economy EU dependence	0.170***	0.179***	0.235**	0.253***	0.253**	0.284***
	(0.027)	(0.026)	(0.100)	(0.085)	(0.098)	(0.077)
Home ownership share	0.299***	0.301***	0.053		0.072	
	(0.042)	(0.042)	(0.197)		(0.201)	
No qualifications share	0.467***	0.535***	0.298		0.393**	0.346**
	(0.070)	(0.079)	(0.184)		(0.171)	(0.147)
No qualifications growth	0.169***	0.183***	0.387***	0.207	0.370***	0.198
	(0.052)	(0.051)	(0.137)	(0.125)	(0.129)	(0.130)
Highly qualified share	-0.643***	-0.563***	-0.115	-0.374***	-0.001	
	(0.091)	(0.102)	(0.290)	(0.124)	(0.285)	
Retail employment share	0.046	0.044	-0.145*		-0.122	
	(0.033)	(0.034)	(0.082)		(0.078)	
Retail employment share change	-0.065**	-0.059**	-0.141*		-0.128*	-0.133
	(0.026)	(0.027)	(0.076)		(0.076)	(0.087)
Construction employment share change	0.049*	0.043	-0.181*	-0.137	-0.136	
	(0.028)	(0.028)	(0.099)	(0.087)	(0.093)	
Finance employment share	-0.058*	-0.034	-0.122*		-0.135**	-0.150**
	(0.031)	(0.034)	(0.069)		(0.063)	(0.060)
Wage change	-0.033		-0.142	-0.196**	-0.168	-0.158
	(0.025)		(0.089)	(0.093)	(0.102)	(0.123)
Interquartile pay range	0.155***	0.183***	-0.087		-0.094	
	(0.045)	(0.046)	(0.171)		(0.166)	
Unemployment rate	0.007		0.430***	0.464***	0.408***	0.440***
	(0.025)		(0.091)	(0.063)	(0.091)	(0.065)
Observations	369	371	95	95	95	95
R2	.849	.844	.661	.628	.675	.652

Notes: Table reports results from OLS regressions. In columns 1 and 2 the dependent variable is the standardized share of the Leave vote in a local authority area in England, Scotland and Wales. In columns 3 and 4 the dependent variable is the standardized share of the Le Pen vote in a French département in the first round of the French presidential election. In columns 5 and 6 the dependent variable is the standardized share of the Le Pen vote in a French département in the second round of the French presidential election. Columns 1, 3 and 5 include the full set of regressors. Columns 2, 4 and 6 show the best subset selection using the AIC information criterion. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A8: Within-City Univariate Analysis of Index of Deprivation Across 107 Wards in 4 Cities

	(1) Leave	(2) Leave	(3) Leave	(4) Leave	(5) Leave	(6) Leave
Index of multiple deprivation: average rank (2015)	4.505***					
	(1.343)					
Income deprivation: average rank (2015)		5.409***				
		(1.348)				
Employment deprivation: average rank (2015)			6.057***			
			(1.264)			
Education and skills deprivation: average rank (2015)				8.259***		
				(1.168)		
Health deprivation: average rank (2015)					5.674***	
					(1.326)	
Crime severity: average rank (2015)						1.805
						(1.255)
Best Subset						
Observations	107	107	107	107	107	107
R2	.203	.241	.274	.414	.24	.137

Notes: Table reports results from OLS regressions. The dependent variable is the Vote Leave share at the ward level across four English cities. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.