Oct 2016 No.305

Who Voted for Brexit?
A Comprehensive District-Level Analysis

Sascha O. Becker, Thiemo Fetzer and Dennis Novy

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

Who Voted for Brexit?
A Comprehensive District-Level Analysis

Sascha O. Becker Thiemo Fetzer Dennis Novy*

October 2016

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 fundamental characteristics of the voting population were key drivers of the Vote Leave share, in particular their age and education profiles as well as the historical importance of manufacturing employment, low income and high unemployment. Migration was relevant only from Eastern European countries, not from older EU states or non-EU countries. We also find an important role for fiscal cuts being associated with Vote Leave. Our results indicate that modest reductions in fiscal cuts could have swayed the referendum outcome. In contrast, even drastic changes in immigration patterns would probably not have made a difference. We confirm the above findings at the much finer level of wards within cities. Our results cast doubt on the notion that short-term campaigning events had a meaningful influence on the vote.

Keywords: Political Economy, Voting, Migration, Austerity, Globalisation, UK, EU

JEL Classification: D72, N44, R23, Z13

Acknowledgments: We thank the ESRC initiative "The UK in a Changing Europe" for financial support. All three authors are based at the University of Warwick and affiliated with the Centre of 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 referendum on European Union membership on 23 June 2016 is thought to have been a watershed moment in European integration and globalisation more broadly. Even though 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.[1] Our paper fills this gap by combining a multitude of geographically disaggregated data sources to carry out a comprehensive analysis of the socio-economic characteristics that correlate with the outcome of the 2016 referendum.

In particular, we study the EU referendum result in England, Wales and Scotland disaggregated across 380 local authorities (and across 107 wards across in cities) and relate this to fundamental socio-economic features of these areas.[2] We focus on socio-economic characteristics that can be broadly grouped into five categories: political variables, 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 as well as 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 obviously multi-causal and multi-faceted. Nevertheless, a systematic analysis across an exhaustive range of socio-economic characteristics can be helpful to direct future research efforts that aim to identify specific

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

[2] 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, the British Overseas Territory, ceded to Britain in 1713 under the Treaty of Utrecht, where many covariates are missing.

mechanisms. One might be able to single out individual predictors 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 prediction performance. In particular, we highlight that the simplest model containing only six explanatory variables capturing electoral preferences as measured by the 2014 European Parliamentary elections explain almost 92 percent of the variation in the support for Leave across local authority areas. This suggests that understanding the evolution of political preferences over time in the UK can provide a unique window into understanding the causal mechanisms that ultimately caused a referendum to be held in the first place.[3]

What socio-economic characteristics do have significant explanatory power for the 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 can be explained by measures of a local authority area's exposure to the European Union (e.g., due to immigration and trade exposure) as well as measures capturing the quality of public services and fiscal consolidation. Rather, a significant amount of the variation can be linked to a range of variables that we can think of as hardly malleable in the short run by political choices (variables such as educational attainment, industry structure and demography). We document that similar patterns hold when exploring data on the EU referendum result across 107 wards in four English cities, which this paper is the first to exploit. Our findings suggest that there is a disconnect between the key correlates of the vote outcome and the topics dominating the political debate in the run-up to the election, in particular the fiscal burden of EU membership and the exposure to European migration.

How can we reconcile this disconnect? The political debate centred on two issues: the fiscal burden of EU membership, which voters may have inadver-

[3]In a panel setup Becker and Fetzer (2016) undertake a first attempt to study the role of immigration from Eastern Europe in shaping a political support base for an anti-EU platform in the UK.

tently evaluated against the sizeable extent of domestic welfare and benefits cuts since the financial crisis, and the exposure to European migration since the enlargement of the European Union in 2004. If we think of migration and fiscal cuts 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 immigration, fiscal cuts and the housing market 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 and a rising age profile).

Migration and austerity were, at least to some extent or at some point in time, political choice variables. We perform a set of back-of-the-envelope calculations based on our estimated models to relate how different choices may have affected the referendum result. The analysis suggests that just a slightly less harsh regime of austerity aimed at cutting benefits could have substantially reduced support for the Vote Leave campaign and overturned the result of the EU referendum. A reduction in migration from Eastern Europe, which could have been achieved by opting to phase in freedom of movement in 2004 (which much of the rest of Europe did), could have also reduced the margin of victory for the Leave campaign, but would have been unlikely to overturn the referendum result.

Lastly, we also explore the role of some short-run factors such as the 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 still have lost on a sunny day.

While our overall results resonate with the anecdotal evidence presented elsewhere, our results point towards two important dimensions that have so far received little attention. First, we show that domestic policy could have counteracted the support for Vote Leave. Our (speculative) back-of-the-envelope calculations suggest that even just a slightly more moderate regime of austerity

could have substantially reduced support for the Vote Leave campaign, which cited questionable figures of the fiscal cost of EU membership.

Second, we highlight that the British first-past-the-post electoral system may have significantly contributed to the emerging chasm in British politics on EU membership. Anti-EU parties, in particular the UK Independence Party (UKIP), have seen strong popular support in European Parliament elections that are based on proportional representation. However, despite significant popular support for UKIP, the party is essentially not represented in the national parliament, implying that a significant share of voters lack formal access to the political system through representation of their views. At the same time, the strong popular support has rightfully attracted media attention. But it has come with no obligation for far-right politicians to assume roles of responsibility towards their electorate by exercising executive power. Instead, political entrepreneurs within the established parties (Labour and especially the Conservative Party) have reached out to UKIP voters in an attempt to strengthen their position within the party, putting significant strain on the internal cohesiveness of their respective parties.

This paper and the Brexit vote it studies can be seen not only in an EU context but is also related to 'populist' campaigning and voting more broadly. A large literature in the social sciences has looked at vote patterns across the political spectrum as a function of demographic, economic and political drivers. The UK, with its first-past-the-post electoral system for the House of Commons, has typically had clear majorities of either the Conservative or the Labour Party since the 1920s. This pattern was broken with the first coalition government in 2010 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 UKIP and the rise of 'nationalist' parties in Scotland, Wales and Northern Ireland. While the latter can be seen as domestic developments driven by a renewed push for devolution (and even independence) of 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. In

Europe more broadly, Arzheimer (2009) analyzes contextual factors explaining far-right voting in Western Europe over the period from 1980 to 2002.

Backlash against globalisation is said to have been another important factor in the Leave vote. 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 saw as an unaccountable force ruling the UK from outside the country. Becker and Fetzer (2016) explore the impact of immigration from Eastern Europe on the support for UKIP in the UK context. Dippel et al. (2015) link votes for far-right parties in Germany to trade-integration with China and Eastern Europe. For the US, Autor et al. (2016) argue that rising trade integration between the U.S. and China contributes to the polarization of U.S. politics. Burgoon (2012) analyzes party opposition to 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, which 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', referenda in other countries are rather rare. The UK has a 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 agreed on 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.

This paper is organised as follows. Section 2 discusses the institutional and historical context of British EU membership. Section 3 presents our empirical approach. Section 4 discusses the underlying data sources and variable groups.

In Section 5 we discuss the main results, while section 6 provides a policy discussion. Section 7 concludes.

2 Britain's role in the EU

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. Ultimately the UK 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 in 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 ref-

erendum, 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 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 most member countries.

To understand the factors that determine EU support, we will exploit a range

of data sources in the empirical analysis that relate the Brexit vote to underlying fundamental characteristics of UK voting areas.

3 Empirical approach

We carry out three main exercises, which we describe here in turn.

3.1 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 turnout T_c or the share of votes to leave L_c.[4]

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

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

Any result from this analysis comes with a grain of salt in that it lacks causal identification. Furthermore, the inclusion of many correlated covariates may make the coefficient estimates unstable and result in high variance of coefficient estimates, which in turn may result in poor out-of-sample prediction performance. For that reason we pursue a model selection approach as our second main approach.

[4]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. Not surprisingly, results are robust to the inclusion of these few additional observations.

3.2 Model selection

In the second approach, we perform a variable selection exercise to identify 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, 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 best subset algorithm 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 quite quickly become infeasible. To see this, suppose that there are k potential regressors. Best subset selection proceeds as follows: the first model estimates – using ordinary least squares (OLS) – all $\binom{k}{1} = k$ 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 $\binom{k}{2}$ models containing exactly two regressors, and so on. In total, $\sum_{p=1}^{k} \binom{k}{p} = 2^k$ models are estimated. With $k = 30$ this amounts to estimating just over one billion regressions. The non-feasibility of best subset selection for large k 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.

3.3 Within-city analysis

While official results are only published at the level of counting areas, we also managed to obtain voting data at the ward level across four UK cities (see section 4.7 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 share within cities. As a matter of fact, the variation within cities is larger than across local authorities.

4 Data

Empirical analysis of election data for the UK is challenging as the data is provided only at a coarse geographic resolution. Data is not provided at the level of individual polling stations. Instead, the EU Referendum Act divided the UK into 382 counting areas (or 'local authority areas'), 327 of which are in England, 22 in Wales and 32 in Scotland. Northern Ireland consists of a single counting area. We generally drop Northern Ireland and Gibraltar so that we end up with 380 local authority areas.[5]

In terms of explanatory variables, we map a multitude of different data sources to the local authority areas. We control for political and socio-economic

[5] A few covariates are missing for some additional local authority areas, which is why some regressions in Tables 1-5 contain fewer observations.

factors to obtain an overarching picture of the correlates of the referendum result. We broadly categorize the variables into five coarse groups of variables: previous voting outcomes, in particular anti-EU sentiment as captured by vote shares for EU-sceptical parties in past European Parliament (EP) elections and in the previous EU referendum of 1975; EU exposure through immigration, trade and structural funds; local public service provision and fiscal consolidation; demography and education; 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 in turn. Finally, 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 use both levels and changes and let the data speak for themselves.

4.1 Voting outcomes

We collect data on turnout and vote shares at the local authority level for the 2016 EU referendum.[6] Vote Leave won 51.9 percent of votes in the EU referendum held on 23 June 2016, 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 of the vote across the local authority areas, while Figure 2 presents the map pertaining to turnout. One striking observation is that some urban centres seemed to have particularly

[6]Corresponding data for the 1975 Referendum were only published for 68 counting areas across the UK. Those roughly followed the historic county borders so that we can use the earlier referendum results across the UK as a covariate capturing historical voting patterns.

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 have affected the overall margin of the result.

The UK parliamentary system with its first-past-the-post voting system makes it difficult to capture shifting political attitudes as voters are less likely to vote for protest parties and get representation in parliament. This explains why the anti-EU UK Independence Party (UKIP) was only ever able to win a single seat in the UK Parliament despite frequently dominating the public debate surrounding the EU.

Elections to the European Parliament, in contrast, rely on a system of proportional representation that helps small parties win seats. In fact, the 2014 European Parliamentary elections saw UKIP win the largest number of British seats. Since those election results are disaggregated by local authority area, we can directly use the UKIP vote share and voter turnout as control variables. The tight correlation between support for UKIP in the 2014 European Parliament elections is presented visually in Figure 3. The positive relationship is striking. A simple regression line has an intercept of around 25 percent and a slope close to unity. 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.[7]

In addition, we also collect data on the 2014 European Parliament vote shares of the far-right racist British National Party (BNP), the Conservative Party, the Labour Party, the Liberal Democrats as well as the Green Party.

[7] Becker and Fetzer (2016) look at election results for European Parliament elections over the period from 1999 to 2014 to study how anti-EU sentiment, in the form of vote shares for the UK Independence Party (UKIP), relates to migration. Their analysis is more restrictive by focusing on migration as the only treatment of interest, but tries to uncover a causal effect of migration. The analysis in this paper is more comprehensive but cannot attempt to address causality for all of the predictors of the referendum vote.

4.2 EU exposure: immigration, trade and structural funds

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, it jumped to 87,000 in 2004. It fell slightly in the aftermath of the global financial crisis when pound sterling also depreciated, only to rise strongly again to an all-time peak of 184,000 in 2015.[8] 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.

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).[9]

The 'take back control' theme of the Leave campaign also extended to the free movement of goods and services. Many voters perceived international trade

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

[9] 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.

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 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.[10] 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).[11]

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 broadly 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). While the gross payment towards the EU budget is not attributable to voting areas, we can track funding received from the EU. In particular, we take account of EU Objective 1 status, which indicates significant transfers for regional development from the EU Regional Policy budget over the years from 1989 to 2013. 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). But on the other hand, EU funding may be perceived by voters as a handout and a symbol of foreign dependence (Davies, 2016).

[10] See Arnorsson and Zoega (2016) for an analysis of the Brexit vote at the level of NUTS2 regions.

[11] We source the data on value added shares from Springford et al. (2016), available here. It combines the contributions of all major sectors to regional GDP (services, manufacturing, construction and primary industries including agriculture, mining and energy supply). Springford et al. (2016) find a *positive* correlation between EU trade integration and the share of voters intending to vote Leave.

4.3 Public service provision and fiscal consolidation

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). 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.

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

We collect data from two main sources capturing the extent of fiscal consolidation and the quality of NHS service delivery across the UK.

First, 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.[12]

Second, 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

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

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.[13] 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.[14]

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 renting on the private market and in council-provided rental housing.

In addition, we use 2011 census data to control for the share of working age residents that commute to Inner London for work. 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.

4.4 Demography and education

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.[15] We capture the education of the local population by the shares of people with various qualification levels.[16]

[13]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.

[14]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.

[15]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 we cannot use an endless number of covariates.

[16]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.

4.5 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. 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.

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.[17] 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.

4.6 Events on the referendum day

Apart from the five groups listed above, 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), 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.[18]

[17] Bell and Machin (2016) report a negative relationship between median wages and the Leave vote share.

[18] The CHIRPS data is available here.

4.7 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 raw 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 behavior. 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 has simply not been collected.[19] 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.[20]

We match this information 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 five deprivation domains. The indices measure Employment, Income, Health, Barriers to Housing, Education Skills and Training Deprivation and Living Environment Deprivation. 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

[19] 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.

[20] 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.

level rank in a specific domain.

The main exercise is a similar variable selection procedure, explaining the variation in the Vote Leave share across wards and cities. Given our list of six ward level characteristics and city fixed effects, best subset selection identifies the underlying location characteristics that drive the percentage of Vote Leave across wards.[21]

5 Results

In section 4, we discussed our variables in different groups. To get a first indication of how these groups are related to the 2016 referendum result, in section 5.1, 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 in a thematic group as predictors of the Vote Leave result. Second, looking at the R^2 for groups of variables informs us about the predictive power of thematic groups relative to each other. In section 5.2 we perform a back-of-the-envelope counterfactual prediction exercise to see how much different important predictor variables would have had to be to overturn the referendum result. After this, in section 5.3, we pool the groups of variables and perform the best subset selection procedure more generally. In section 5.4 we extend our analysis to more finely disaggregated geographic areas by exploiting within-city variation across wards of four UK cities. Finally, in section 5.5, 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 bad fundamentals.

5.1 Predicting the Brexit vote by variable group

All of the five tables pertaining to results for the five groups of predictor variables 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

[21] 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.

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 3 is a case in point. For this reason, in Tables 1 to 5 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.

Group 1: 1975 Referendum and 2004 European Parliament Elections We start by looking at vote shares in previous elections as predictors for the EU Referendum result. This exercise is slightly different from the subsequent tables in that one might think of referendum results as a function of (linear combinations of) lagged values of previous election results. It is therefore hardly surprising to find high predictive power, even when only using the UKIP vote share in the 2014 European Parliament (EP) elections as the sole predictor in column 1. The variation from this one variable alone generates an R^2 of 0.751.[22] Adding the vote share of the British National Party 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 0.874 of the variation in the referendum result. The subsequent columns add only marginally to the already large R^2. Overall, the full set of regressors explains 0.919 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.

Looking at signs and point estimates, Table 1 shows that both the UKIP vote share and the BNP vote share are positively correlated with the Vote Leave share. The interpretation of coefficient point estimates is simplified by the fact

[22] This regression corresponds to Figure 3.

that all continuous variables 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 UKIP vote share in the 2014 EP elections is associated with a 7.51 percentage-point larger Vote Leave share. The vote shares for the Conservative Party, the Labour Party and the Green Party all have a negative sign, and the Liberal Democrat vote share is statistically insignificant.

Overall, Table 1 suggests that understanding the UKIP vote share is crucial for understanding the vote in favour of Brexit. Only founded in 1991 and taking on its current name in 1993, UKIP is a fairly new contestant on the British political scene. It has been traditionally seen as pushing the single issue of Britain leaving the European Union. 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. But due to Britain's first-past-the-post voting system UKIP is otherwise hardly represented in national UK politics. UKIP only has a single Member of Parliament, zero representation on parliamentary select committees in the House of Commons and only three representatives in the House of Lords.[23] Understanding the drivers behind UKIP's ascent in recent years thus seems important (Ford and Goodwin, 2014).

Interestingly, the Leave share in the 1975 EU Referendum has no predictive power whatsoever. The arguments for and against EU membership were quite different from those in the 2016. In fact, over 67 percent of voters supported the Labour government's campaign to stay in the EEC, or Common Market, despite several cabinet ministers having come out in favour of British withdrawal.[24]

Group 2: EU Exposure (Immigration, Trade and Structural Funds) In Table 2 we correlate the Vote Leave share with measures of immigration, EU trade dependence and EU subsidies (Structural Funds). It turns out migrant shares (from the 12 EU accession countries that joined in 2004 and 2007, the initial EU 15 countries, and non-EU countries) from the year 2001 in *levels* are negatively

[23]This is correct as of late 2016.

[24]The fact that results for the 1975 EU Referendum were only published at the level of 68 counting areas arguably reduces the predictive power.

correlated with the Brexit vote as those immigrants predominantly moved to urban areas that in 2016 voted for Remain. The striking observation is that in terms of migration *growth*, only migration from the mainly Eastern European EU accession countries *positively* correlates with the Vote Leave share. The EU trade dependence of local authority areas is also positively correlated with the Vote Leave share as areas with a heavy concentration of manufacturing (such as the North East of England), which tend to disproportionately import from and export to European Union countries, were likely to Vote Leave. Finally, 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. Overall, variables capturing EU exposure explain just below 50 percent of the variation in the Vote Leave share.

Group 3: Public Service Provision and Fiscal Consolidation In Table 3, we observe that the share of residents in a local authority area who commute to London is a strong predictor for voting Remain.[25] This might be explained by the fact that those commuting into London are relatively high-skilled who had a larger tendency to vote Remain. On the other hand, house ownership is strongly correlated with the Vote Leave share. This correlation may be not be surprising as house ownership is highest amongst the older section of the population. The share of the population in council rented 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

[25]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.

more likely to vote in favour of leaving the EU. 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.

Group 4: Demography and Education In Table 4, we explore whether demography and education variables predict the Referendum result. As predictors, we use both the levels in 2001 and the growth between 2001 and 2011 of the share of the population that has no qualification or a high qualification, respectively. The middle qualification range is the reference group. Results indicate that a larger baseline share of the population with no qualification is associated with a larger Vote Leave share. An stronger increase in that share between 2001 and 2011 further increases the Vote Leave share. In contrast, the share of the population that has a high qualification is associated with a lower Vote Leave share. A faster growth of the share with no qualification is also associated with a larger Vote Leave share. Somewhat surprisingly, the same is true for a faster growth in the share of those with a high qualification. 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. Only the analysis on the full set of variables in Section 5.3 will allow us to see this more clearly.

In terms of age brackets, we use the share of the population of age 60 and older, which makes those younger than 60 the reference group.[26] Both a higher baseline level 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 that there was a clear age gradient in the Vote

[26]Note that, in principle, we could use more finely grained age brackets, but in the long specifications in Section 5.3, this runs into dimensionality issues for the machine learning algorithm, as explained above.

Leave share, with younger voters intending to vote Remain and older voters Leave. It is striking that this group of variables has the largest predictive power of any of the groups, except for the election variables in Table 1, with an R^2 of up to 78 percent.

Group 5: Economic Structure, Wages and Unemployment In Table 5, we concentrate on variables characterizing the sectoral structure and median hourly pay, both in levels (in the base year 2001) and their change over the decade 2001 to 2011, as well as unemployment rate, self-employment rate and participation rate, in the year prior to the referendum. After some experimentation, 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, public sectors and various service sectors. This being said, a higher share of employment in the base year in any of the four sectors highlighted in Table 5 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 retail employment (arguable a mix between blue- and white-collar jobs) is associated with a lower Vote Leave share, whereas a stronger increase in manufacturing employment is associated with a higher Vote Leave share, whereas the growth of employment in construction or finance is not significantly associated with the Vote Leave share. A higher median hourly pay in the base year 2005 as well as a stronger increase in it are both associated with a lower Vote Leave share, consistent with the narrative that those "left behind" were more likely to Vote Leave. Finally, a larger unemployment rate is associated with a larger Vote Leave share, but the self-employment rate and the overall labour market participation rate have no predictive power for the Vote Leave share. Overall, variables in this group explain nearly 68 percent of the variation in the Vote Leave share.

Summary of analysis of five groups of predictor variables Overall, each of Tables 1-5 yields an R^2 of at least 46 percent with a full set of regressors. However, the strongest explanatory power lies with political variables in Table 1 as

well as demography and education variables in Table 4. Figure 5 gives a visual overview of the goodness of fit across Tables 1-5.

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 5.3.

5.2 Counterfactual scenarios

The results in Tables 2-5 lend themselves to back-of-the-envelope counterfactual scenarios. 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 consider a few key variables and calculate the counterfactual differences in these variables that would have led to the opposite referendum outcome. We highlight that these scenarios are somewhat speculative for various reasons, as we explain further below.

We start with Table 2 and regression results on migrant growth in column 8. While migrant growth from EU 15 countries between 2001 and 2011 carries a negative coefficient, the growth of migrants from EU accession countries is the only growth rate positively (and significantly) related to the Vote Leave share. The corresponding coefficient stands at 1.430, meaning that *ceteris paribus* a one standard-deviation lower migrant growth rate would have been associated with a 1.430 percentage point lower Vote Leave share. The migrant growth rate from EU accession countries would therefore have had to be $1.329 = 1.9/1.430$ standard deviations lower to make a difference to the referendum outcome. Taken literally, this means that instead of the 1.7 percent growth rate as actually observed between 2001 and 2011 across local authority areas on average, the UK would have had to experience a -0.5 percent growth rate, i.e., a decline in immigrants from EU accession countries.[27] Therefore, only a dramatic reversal

[27]The average increase in the share of migrants from EU accession countries is 0.0170 percent with a standard deviation of of 0.0169. Given that the standardized growth rate would have had to be 1.329 standard deviations lower, the counterfactual growth rate follows as $-0.005 = 0.017 - 1.329^{*}0.0169$. Note that the corresponding coefficients in columns 5-7 of Table 2 are higher (although not significantly different from the value used above) so that the required shift

of the EU accession immigration experience would have swayed the vote.

We stress that our counterfactual scenario is based on a linear projection. That is, we calculate the same across-the-board counterfactual change in the migrant growth rate across local authority areas, holding all other variables constant. We follow this linear projection to retain comparability to further counterfactual scenarios below.[28]

We then consider another key policy variable – the role of fiscal cuts in Table 3. The corresponding coefficient in column 8 stands at 5.637, meaning that fiscal cuts would have had to be 'only' $0.337 = 1.9/5.637$ standard deviations lower to overturn the referendum result. Given that the average fiscal cut was £448 per person, this translates into counterfactual cuts that would have had to be £41 lower per person.[29] This reduction corresponds to barely 10 percent of total fiscal cuts (and *not* to 10 percent of total fiscal spending). Such a policy change would therefore have appeared eminently feasible compared to the drastic reduction in migration described above. Given the current UK population of around 65 million, the aggregate reduction in fiscal cuts would have amounted to roughly £2.7 billion.[30] 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.

We stress that the £41 result should be interpreted as a lower bound in the sense that larger reductions in cuts may have been required. 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.

Moving on to column 6 of Table 4, we focus on the share of the population with no qualifications as well as the share of the population aged 60 and older, both in the base year. The corresponding coefficients are 7.189 and 0.813. The

in the growth rate would have had to be somewhat weaker.

[28]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.

[29]The standard deviation of fiscal cuts in the data is £122. The counterfactual level of cuts follows as £407 = £448 − 0.337*£122.

[30]This compares to a net UK contribution to the EU budget of around £8.5 billion in the year 2015.

counterfactual predictions, separate for each variable, imply that the share of the population with no qualifications would have had to be 33.6 percent instead of 35.4 percent, while the share of the population aged 60 and older would have had to be 12.3 percent instead of 24 percent.[31] 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 to differences in senior citizen population shares.

Finally, we consider column 13 in Table 5. We are interested in the manufacturing employment share in the base year (with a coefficient of 4.117), median hourly pay in the year 2005 (with a coefficient of -1.412) and the unemployment rate in 2015 (with a coefficient of 0.860).[32] The counterfactual predictions, again separate for each variable, yield the following results. The manufacturing employment share would have had to be 12.6 percent instead of 15.1 percent. Median hourly pay would have had to be £13.66 instead of £10.97. The unemployment rate would have had to be 0.6 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 pay and unemployment.

Overall, such counterfactual scenarios must be interpreted with a grain of salt of course, not least since various regressors on the right-hand side are correlated. There are no immediate policy implications because it is not clear that, say, an across-the-board upgrading of qualifications across the population would change the relative distribution of economic gains and losses from EU membership, let alone political sentiment. Nevertheless, we get a sense of the relative importance of the various regressors when comparing the threshold standard deviations necessary to push the Vote Leave share under 50 percent.

5.3 Best subset selection results

In Table 6 we use the best subset selection procedure for variables across all groups. Column 1 displays the best subset of variables when all variables from

[31] 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.

[32] The corresponding standard deviations of the non-standardized variables are 5.4 percent, £2.00 and 2.1 percent, respectively.

the previous five separate groups of variables are combined in one joint 'horse race.' In column 1, five political variables are included, e.g., the vote shares in the 2014 European Parliament elections for UKIP, BNP, the Conservative Party, the Labour Party, the Green Party, and the Liberal Democrats, as well as ten non-political regressors. Overall, we obtain an R^2 of over 95 percent with just 15 variables.

In column 2, we remove all vote share variables – all of which could be considered, with a grain of salt, as lagged dependent variables – but still obtain an R^2 of 87 percent. The reason is that the substituted regressors such as the qualification profile of the population are highly correlated with the political variables. This suggests that the political variables are driven by the same fundamentals as the Vote Leave share.

For comparison, column 3 displays a full specification including all variables. Columns 4 through 8 re-display estimates using only the best subsets uncovered in each of the five groups of variables from previous tables.

Overall, while political variables are clearly important, they can largely be explained by underlying socio-economic fundamentals. The overlap between columns 1 and 2 in terms of regressors is minuscule, yet column 2 yields a similar R^2. Still, it is hard to discern the underlying mechanism or relationship between political and socio-economic fundamentals because Table 6 just reports conditional correlations with no causal identification. The signs are especially hard to interpret because the regressors are correlated with each other. But the fact that the bulk of variables capture the local economic structure (transport infrastructure, qualifications, education and their changes) tells us that in their entirety they are important.

There was a change in voters' attitudes, as evidenced by the growing UKIP vote shares over time.[33] Using credible identification strategies and possibly micro-level data on individual voters is an important future research agenda to explain the patterns in voter migration in response to changes in socio-economic fundamentals. Nevertheless, the next section 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

[33]See Becker and Fetzer (2016) for details on the rise of UKIP.

are not just masking an urban vs. rural divide.

5.4 Exploiting within-city variation

We perform an analysis of the EU referendum results 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 studying finer spatial variation, or whether the fact that support for the Remain side was more pronounced in cities is driven by compositional 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, are masking rural versus urban differences. We will document here that, when studying within-city variation of the support for the Leave side in the referendum, we observe very similar gradients in proxies for fundamentals, such as educational attainment, suggesting that the fact that a lot of cities appear to have supported the remain side is simply driven by a compositional effect.

Unfortunately, at the ward level, we do not have the same breadth of variables available. 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.[34] We convert the data to standardize the average rank into z-scores, indicating that a higher rank means a higher degree of deprivation, such that when we relate the indices of deprivation with the Vote Leave share in a ward, a positive coefficients is implies that worse fundamentals are associated with a higher Leave share.

Since the sub-indices of the overall deprivation index are very highly cor-

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

related 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 regression between the Vote Leave share at the ward level and the overall index as well as various sub-indices. A best subset selection exercise akin to the previous analysis across local authority areas suggests that we should include city fixed effects (the results for this are presented in Appendix Table A7). Hence, we include city fixed effects throughout such that all the residual variation stems from variation across wards within cities.

The results from this exercise are presented in Table 7. All across the board, more deprivation is associated with a larger Vote Leave share or, vice versa, less deprivation is associated with a lower Vote Leave share. The important point to observe here again is that the tightest relationship between the support for the Leave side is stemming from the sub-index capturing deprivation in education and skills. This is further highlighted in Figure 7, which plots the univariate correlation between the deprivation in Education, Skills and Training and the support for Vote Leave after city fixed effects have been removed.

The results here suggest that similar demographic and socio-economic forces were driving the support for the Leave side within as well as across cities and do not mask a mere rural-urban divide.

5.5 Interaction terms

While we have so far concentrated on a comprehensive approach to predicting the Vote Leave share, we also want to highlight whether the interaction of 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, 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 of new migrants from East-

ern 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 as "flow" variables in separate regression analyses.

Each of these flow variables are being interacted with one of the following "stock" variables: the share of the workforce with no qualification; the sectoral share of manufacturing; the sectoral share of finance, all three measured in 2001; the median hourly pay in 2005. The results are interesting and highlight some important aspects. The main effect of the stock variables that characterize "pre-existing conditions" in the first row of Table 8 are consistent across all four different "flow" variables: the share of the population with no qualification 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 are 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. But the interaction terms, which are the main focus of interest in this exercise show an interesting 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. Interestingly, 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 smaller and often insignificant. This suggests that migration from Eastern Europe, which was different in nature, namely 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 bad fundamentals, such as a workforce with low qualifications, this implies that the incidence of the cuts in per capita terms is strongly correlated with these bad fundamentals. In fact, the correlation between the share of households with no qualification and the Total fiscal cuts measure is around 0.7, 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.

5.6 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. Appendix Tables A1 through A6 present those results. We focus here, for the sake of brevity, just on the discussion of Table A1: local authority areas with larger vote shares in the 2014 European Parliament elections for UKIP, the Conservative or Liberal Democrats had significantly higher turnout in the 2016 referendum, whereas areas with a larger Labour vote share had lower turnout, consistent with the criticism raised against Labour leader Jeremy Corbyn for fighting a lukewarm campaign that did not mobilise Labour voters.

6 Further discussion

6.1 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 since the 1990s, the UK Independence Party (UKIP), had received dramatic political support in European Parliamentary elections. In the 2014 election British voters made UKIP

the largest delegation to the European Parliament. At the same time, UKIP is essentially not represented domestically in the House of Commons, currently only having one Member of Parliament (see our discussion in section 4.1). This chasm between popular support and domestic political representation is a result of the first-past-the-post electoral system which has greatly favoured the established parties, in particular Labour and the Conservatives.

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 the party and its wider platform have not been subject to the same level of scrutiny and oversight by the media. At the same time, however, the media recognized that UKIP had a popular mandate by the electorate. Led by the tabloid press, they provided UKIP with ample coverage but without the corresponding scrutiny, partly because UKIP politicians never had to deliver political outcomes at the national level and were therefore difficult to hold responsible.

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 stewardship.[35]

[35]Italy's populist 5-star movement (M5S) seems to be facing a reality check now that party

In our view, it is 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.

6.2 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).[36] 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 section 2). Margaret Thatcher negotiated the UK budget rebate in 1984. The UK opted out of the 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, ed-

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).

[36]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.

ucation 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 some parts of the economy can be observed in other EU countries. 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.

6.3 Disputable short-run narratives

We are unconvinced by narratives that explain the Brexit vote based on short-run phenomena. First, as we show above, rain and train cancellations on the day of the vote are not systematically related to the voting outcome. Second, post-referendum polls indicate that most people would vote in the same way even after having observed the precipitous fall of sterling in the immediate aftermath of the vote. 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 be felt by abstainers (Curtice, 2016).

We are also sceptical of narratives suggesting that 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 do we believe that the murder of Jo Cox, a Member of Parliament from Yorkshire, in the week before the referendum shifted the vote in any major way.

Instead, our analysis suggests that fundamentals about the economy and characteristics of the population predict the voting outcome reasonably well. The same fundamentals also explain the vote share for the UK Independence Party (UKIP) more generally, suggesting that similar factors are at work.[37]

[37] In fact, regressing the Leave support in the Brexit referendum on a constant and the UKIP support at the 2014 European Parliament elections yields an R^2 of over 70 percent, see 1 and Goodwin and Heath (2016).

6.4 The role of non-economic factors

Those fundamentals 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. 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.[38] Based on detailed Gallup survey data on individuals, Rothwell (2016) makes similar observations about supporters of Donald Trump's presidential campaign in the U.S. He 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.

Overall, a more complex picture arises about the challenges of adapting to social and economic change. In addition, it is likely that people self-select into local areas that better fit their outlook on life, for instance socially liberal and fairly affluent young professionals concentrating in the London area. This self-selection might explain the perceived increase in political polarization between 'cosmopolitan' and 'provincial' areas (Jennings and Stoker, 2016).

6.5 Betting markets and polls

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 North-west 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 Re-

[38] 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."

[39] Burnham warned of "too much Hampstead and not enough Hull." See here and here.

main campaign.[40]

Likewise, 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."

Overall, we can only speculate that large parts of the media and the 'Westminster bubble' were out of touch with the sentiment felt by the majority of voters, in particular in England and Wales. Perhaps the problem was compounded by 'group-think' pressure on pollsters to adjust their figures to the average. These observations are consistent with the 'contact hypothesis' in that too many key actors in public life are concentrated amongst their own kind and lack interaction with outsiders so that they fail to understand their concerns.

7 Summary and policy conclusions

In summary, we find that the 2016 Brexit referendum result is tightly correlated with previous election results for the UK Independence Party, and also those of the extreme right British National Party. Both are related to various fundamental characteristics of the voters across the 380 local authority areas we consider. In particular, as widely reported in the press, the share of the population aged 60 and above as well as the share of the population with little or no qualifications are strong predictors of the Vote Leave share. Furthermore, areas with a strong tradition of manufacturing employment were more likely to Vote Leave, and also those areas with relatively low pay and high unemployment.

We also find strong evidence that the growth rate of migrants from the 12 EU accession countries that joined the EU in 2004 and 2007 is tightly linked to

[40]See Peter Mandelson's account of the Remain campaign here and also here.

the Vote Leave share. This stands in contrast to migrant growth from the EU 15 countries or elsewhere in the world. We therefore conclude that migration from predominantly Eastern European countries has had a distinct effect on voters. However, we cannot identify the precise mechanism – whether the effect on voters is mainly economic through competition in the labour and housing markets, or rather in terms of changing social conditions.

In addition, we find that the quality of public service provision is also systematically related to the Vote Leave share. In particular, fiscal cuts in the context of the recent UK austerity programme are strongly associated with a higher Vote Leave share. We also produce evidence that lower-quality service provision in the National Health Service is associated with the success of Vote Leave.

We are also able to obtain voting data at the level of wards across four cities in the UK. Those are areas with only 7,000 residents on average. We confirm that even across wards within cities (for instance, across wards in the Borough of Greenwich in London), 'weak' socio-economic fundamentals are strong predictors of the Vote Leave share. The same pattern as across local authority areas can thus be observed across wards, even within large cities. The fact that support for Vote Leave may be less visible in a large city like London is merely down to composition effects in that London has relatively strong socio-economic fundamentals on average compared to the rest of the country.

Our regressions allow for a counterfactual analysis. We find that relatively modest reductions in fiscal cuts at the local authority level (less than £50 per person) may have been sufficient to lead to the opposite referendum outcome, pushing the Vote Leave share below 50 percent. The overall reductions in fiscal cuts would have amounted to less than £3 billion in total for the UK. In contrast, even major changes to immigration from Eastern Europe would have been very unlikely to sway the vote in any meaningful way.

In terms of policy conclusions, we argue that the voting outcome of the referendum 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 older, less educated and confronted with below-average public services. We therefore doubt that a different style of short-run campaigning would have made a meaningful difference

to vote shares. Instead, a more complex picture arises about the challenges of adapting to social and economic change. In addition, it is likely that people self-select into local areas that better fit their outlook on life, for instance socially liberal and fairly affluent young professionals concentrating in the London area.

It is clear that a majority of politicians and the media were caught off guard by the referendum result. This speaks to the polarization between metropolitan and other areas. We argue that the 'Westminster bubble' is key to understanding the voting outcome. The under-representation of anti-EU parties in the British parliament is likely a crucial contributing factor to the lack of attention paid in the political process to struggling areas, especially in England and Wales. As a result of the first-past-the-post voting system, UKIP currently only has one 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 introducing more proportional representation into British politics.

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.

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

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

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

Burn-Murdoch, J. (2016). 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.

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*.

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

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

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.

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.

Guyon, I. and A. Elisseeff (2003). An Introduction to Variable and Feature Selection Isabelle. *Journal of Machine Learning Research 52*(3), 239–281.

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.

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*.

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.

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. (2016). Explaining Nationalist Political Views: The Case of Donald Trump. *Gallup Working Paper, August 2016*.

Springford, J., P. McCann, B. Los, and M. Thissen (2016). Brexiting Yourself in the Foot: Why Britain's Eurosceptic Regions Have Most to Lose from EU Withdrawal. *Centre for European Reform, London*.

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

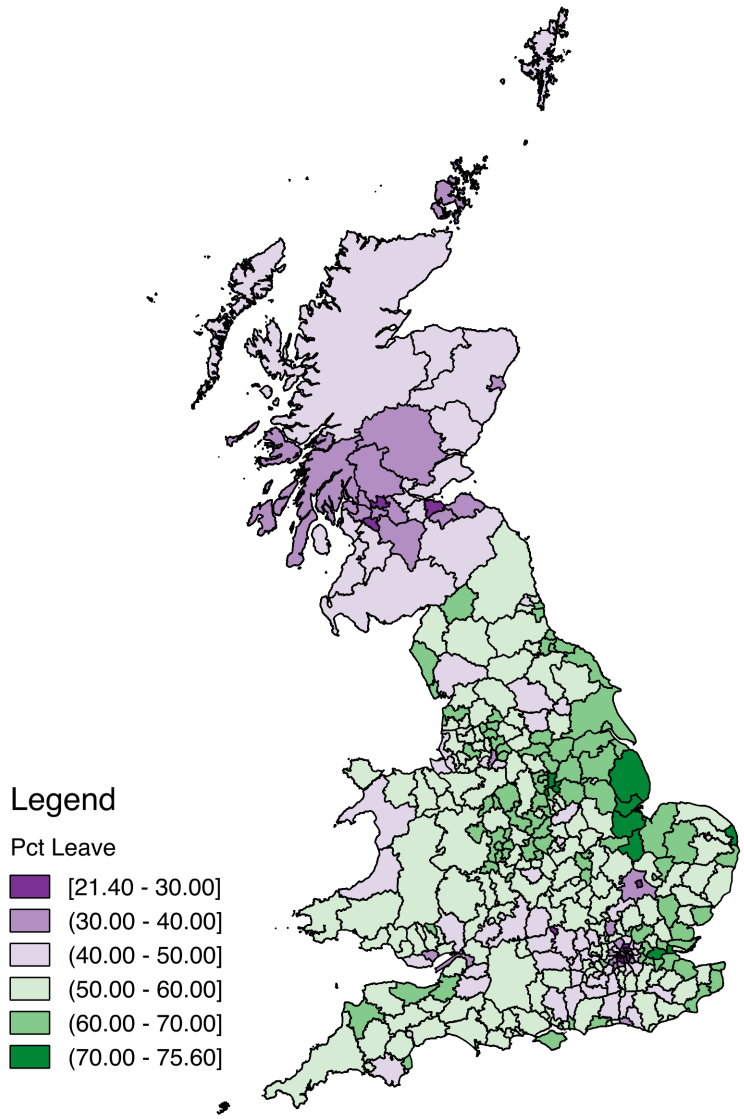

Figure 1: Map of the Leave share (in percent) across local authority areas in the 2016 EU referendum.

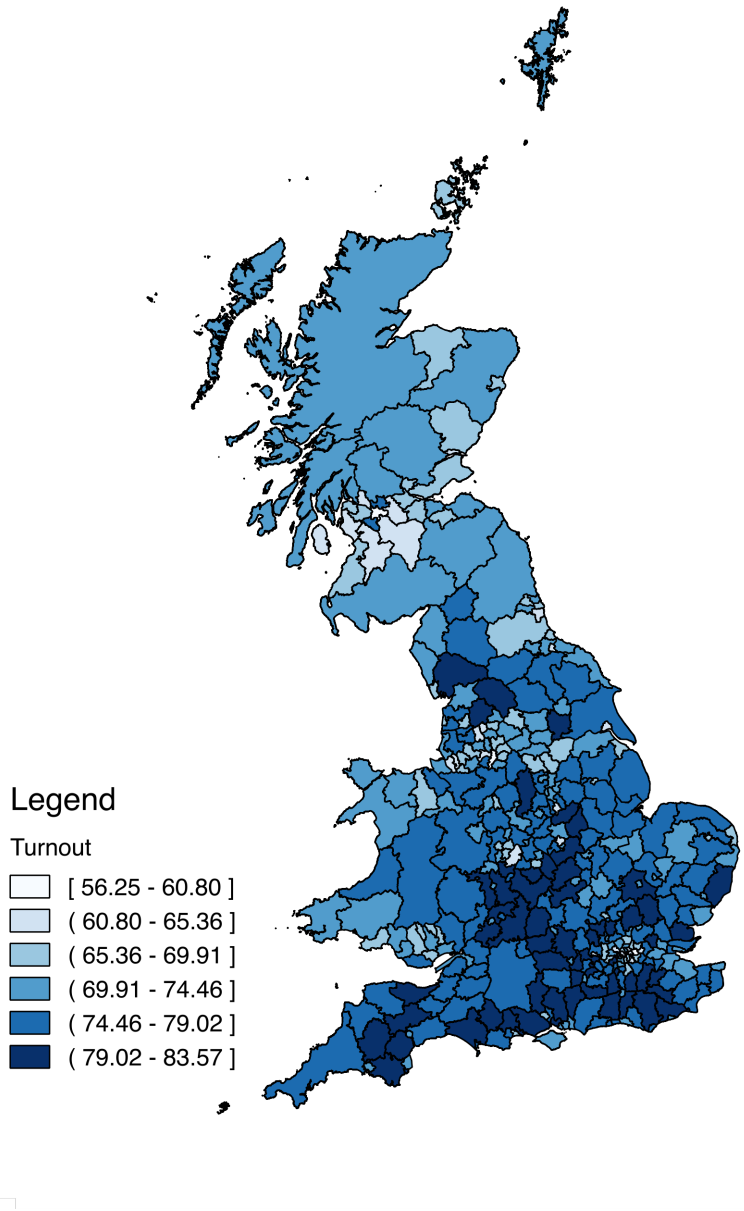

Figure 2: Map of turnout (in percent) across local authority areas in the 2016 EU referendum.

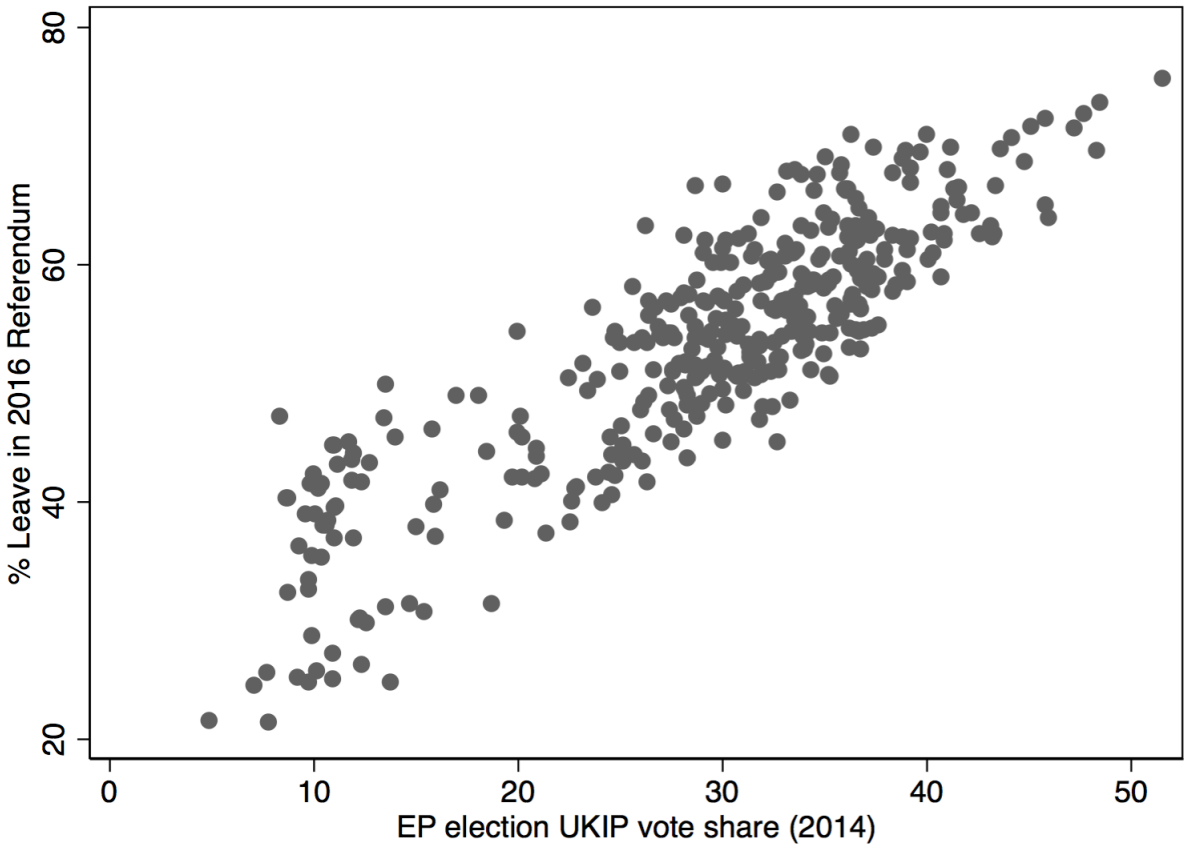

Figure 3: The figure shows the UKIP vote shares (in percent) across local authority areas in the 2014 European Parliament elections plotted against the Leave shares in the 2016 EU referendum.

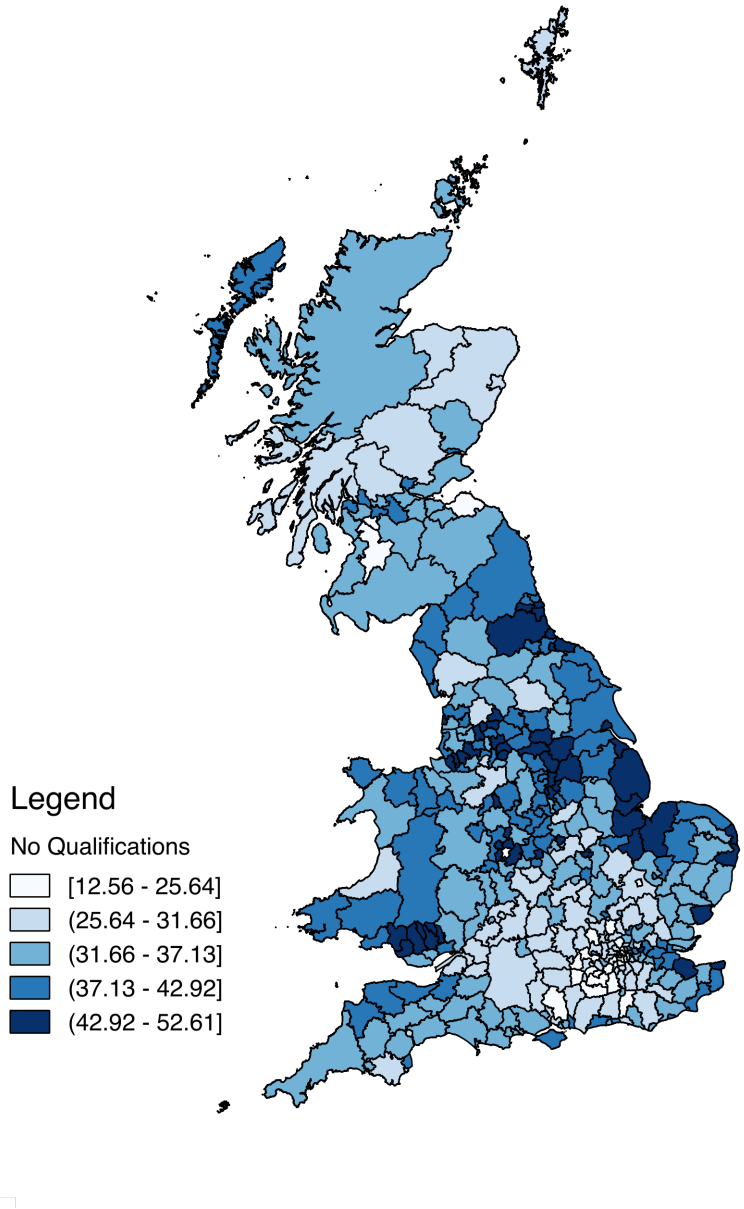

Figure 4: Map of the share of 16-74-year-olds with no qualifications in the year 2001 across local authority areas.

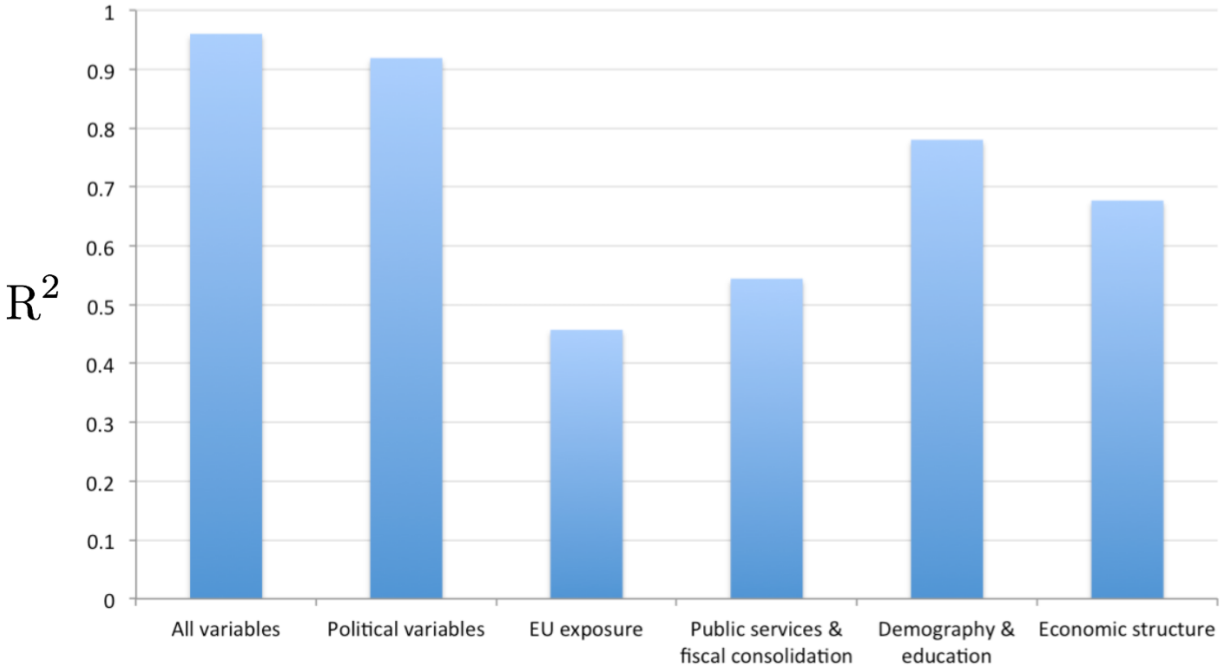

Figure 5: Goodness of fit (measured as R^2) in separate regressions explaining the Leave vote share at the local authority area level using only regressors from the respective group of variables.

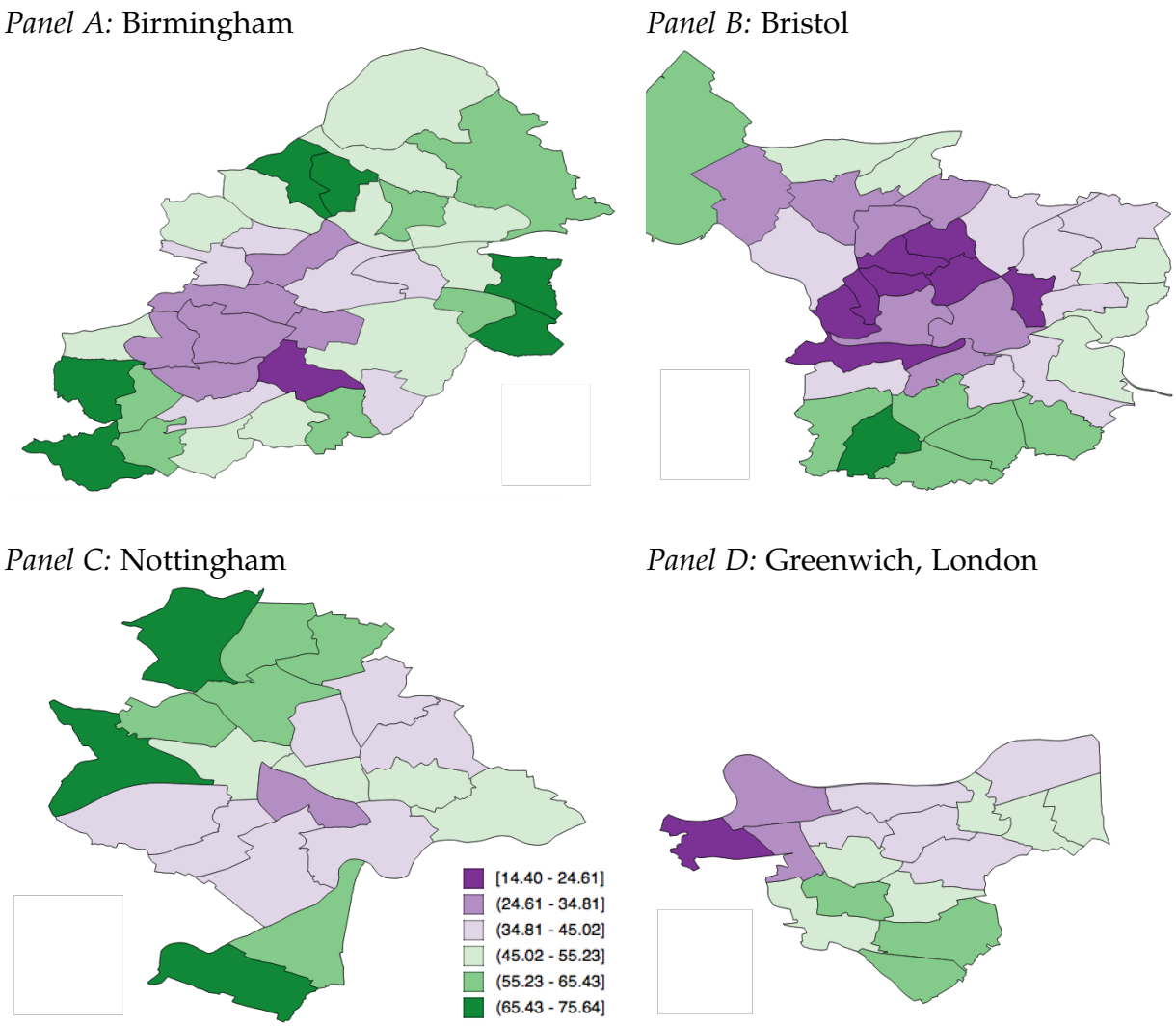

Figure 6: Map of the Leave share (in percent) across wards in Birmingham, Bristol, Nottingham and the Royal Borough of Greenwich in London in the 2016 EU referendum.

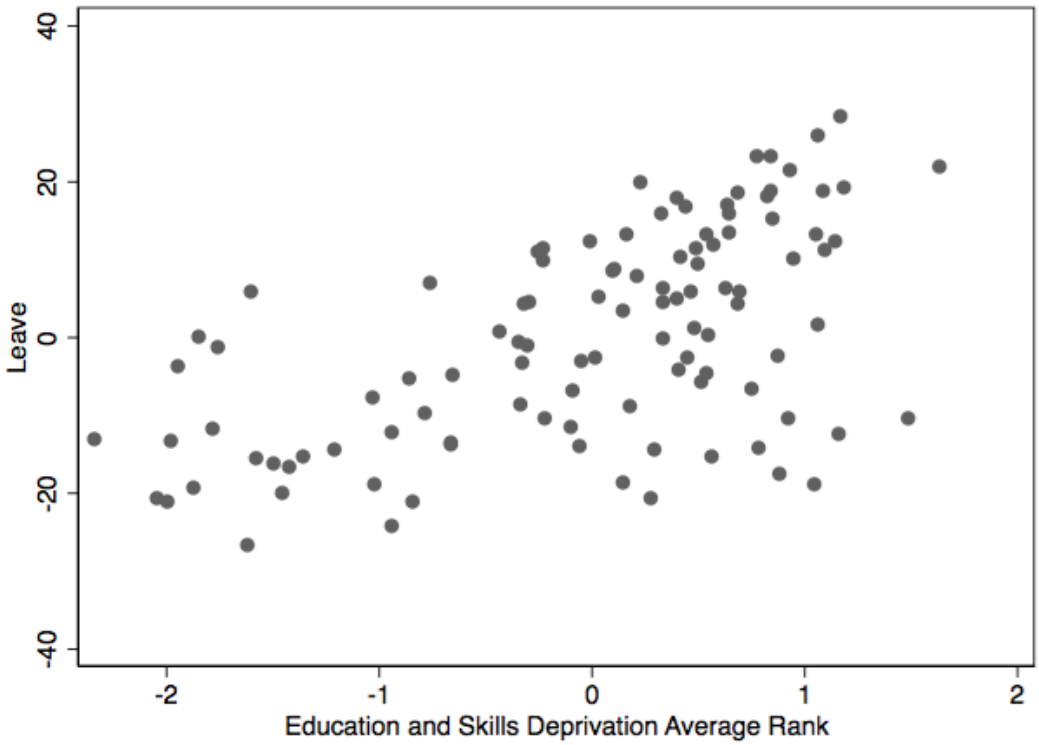

Figure 7: Relationship between the Education, Skills and Training Deprivation Rank and the Leave share (in percent) exploiting variation across 107 wards within four cities (Birmingham, Bristol, Nottingham and Greenwich/London).

Table 1: Predictors of Brexit Vote: 1975 Referendum and 2014 European Parliament Elections

	Benchmark							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EP election UKIP vote share (2014)	9.027***	8.106***	7.613***	8.078***	7.682***	7.514***	7.573***	7.581***
	(0.306)	(0.261)	(0.213)	(0.245)	(0.262)	(0.294)	(0.347)	(0.346)
EP election British National Party vote share (2014)		3.765***	2.683***	2.153***	2.492***	2.544***	2.537***	2.549***
		(0.206)	(0.193)	(0.197)	(0.246)	(0.257)	(0.256)	(0.253)
EP election Conservative Party vote share (2014)				-1.007***	-1.588***	-1.694***	-1.627***	-1.651***
				(0.228)	(0.316)	(0.341)	(0.407)	(0.417)
EP election Labour Party vote share (2014)					-1.102***	-1.413***	-1.364**	-1.379***
					(0.413)	(0.502)	(0.528)	(0.527)
EP election Green Party vote share (2014)			-2.342***	-2.330***	-2.390***	-2.283***	-2.257***	-2.265***
			(0.260)	(0.252)	(0.253)	(0.258)	(0.276)	(0.283)
EP election Lib-Dem vote share (2014)						-0.369	-0.362	-0.372
						(0.261)	(0.264)	(0.272)
EP election turnout (2014)								0.056
								(0.209)
1975 Referendum Leave share							0.094	0.100
							(0.297)	(0.297)
Best subset						X		
Observations	380	380	380	380	380	380	380	380
R^2	.751	.874	.909	.915	.918	.919	.919	.919

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: EU Exposure (Immigration, Trade and Structural Funds)

	Benchmark							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Initial EU accession migrant share (2001)				-1.122	-1.793***	-1.376	-1.427	-1.429
				(0.793)	(0.637)	(0.891)	(0.891)	(0.896)
EU accession migrant growth (2001-2011)					1.623***	1.929***	1.419**	1.430**
					(0.570)	(0.591)	(0.656)	(0.689)
Initial EU 15 migrant share (2001)	-5.665***	-4.739***	-3.884***	-3.307***	-2.945**	-2.806**	-2.930**	-2.935**
	(0.893)	(0.854)	(1.024)	(1.152)	(1.164)	(1.180)	(1.191)	(1.177)
EU 15 migrant growth (2001-2011)			-1.266*	-1.013	-1.609**	-1.167	-1.318*	-1.315*
			(0.699)	(0.698)	(0.676)	(0.764)	(0.781)	(0.767)
Initial non-EU migrant share (2001)						-1.224	-0.661	-0.635
						(0.950)	(0.991)	(1.246)
Non-EU migrant growth (2001-2011)								-0.039
								(0.859)
Total economy EU dependence		3.896***	3.629***	3.547***	3.460***	3.415***	3.288***	3.286***
		(0.407)	(0.417)	(0.407)	(0.405)	(0.408)	(0.516)	(0.521)
EU Structural Funds per capita (2013)							-0.141	-0.143
							(0.517)	(0.517)
Best subset					X			
Observations	380	380	380	380	380	380	369	369
R^2	.296	.428	.434	.44	.457	.46	.461	.461

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: Public Service Provision and Fiscal Consolidation

	Benchmark							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share 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		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 suspected cancer patients 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
R^2	.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 4: Predictors of Brexit Vote: Demography and Education

	Benchmark					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of pop. no Qualification (2001)				4.671***	6.649***	6.608***
				(0.719)	(0.884)	(0.886)
Share of pop. no qualification growth (2001-2011)		3.164***	3.318***	4.293***	5.355***	5.371***
		(0.504)	(0.467)	(0.549)	(0.582)	(0.578)
Share of pop. qualification 4+ (2001)	-8.208***	-10.342***	-10.178***	-6.706***	-6.106***	-6.074***
	(0.434)	(0.590)	(0.564)	(0.782)	(0.827)	(0.818)
Share of pop. qualification 4+ growth (2001-2011)					2.139***	2.110***
					(0.473)	(0.476)
Share of pop. 60 older (2011)						0.308
						(0.369)
Share of pop. 60 older growth (2001-2011)			1.866***	2.805***	2.466***	2.267***
			(0.269)	(0.292)	(0.290)	(0.335)
Best subset					X	
Observations	380	380	380	380	380	380
R^2	.621	.671	.703	.745	.771	.771

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: Economic Structure, Wages and Unemployment

	Benchmark												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Retail employment share (2001)	7.019***	5.514***	4.302***	4.186***	4.381***	4.233***	4.260***	3.872***	3.938***	3.876***	3.857***	3.858***	3.857***
	(0.418)	(0.403)	(0.430)	(0.407)	(0.393)	(0.392)	(0.384)	(0.420)	(0.414)	(0.409)	(0.416)	(0.416)	(0.424)
Retail employment share change (2001-2011)				-1.443***	-1.596***	-1.469***	-1.230***	-1.661***	-1.389***	-1.507***	-1.464***	-1.454***	-1.457***
				(0.371)	(0.377)	(0.362)	(0.345)	(0.396)	(0.385)	(0.430)	(0.452)	(0.452)	(0.468)
Manufacturing employment share (2001)		3.621***	3.688***	3.767***	3.679***	3.738***	3.930***	3.721***	4.012***	4.120***	4.124***	4.122***	4.117***
		(0.356)	(0.302)	(0.290)	(0.292)	(0.292)	(0.299)	(0.381)	(0.403)	(0.423)	(0.423)	(0.423)	(0.447)
Manufacturing employment share change (2001-2011)							0.760**		0.818**	0.999*	1.010**	1.003**	1.005*
							(0.369)		(0.392)	(0.510)	(0.503)	(0.505)	(0.513)
Construction employment share (2001)			3.220***	2.962***	2.974***	3.264***	3.182***	3.037***	2.977***	2.964***	2.923***	2.915***	2.915***
			(0.426)	(0.418)	(0.409)	(0.428)	(0.430)	(0.461)	(0.457)	(0.460)	(0.496)	(0.512)	(0.513)
Construction employment share change (2001-2011)										0.374	0.364	0.361	0.363
										(0.477)	(0.481)	(0.480)	(0.483)
Finance employment share (2001)								0.942**	1.088**	1.081**	1.100**	1.106**	1.106**
								(0.472)	(0.464)	(0.463)	(0.476)	(0.482)	(0.485)
Finance employment share change (2001-2011)											-0.115	-0.114	-0.113
											(0.392)	(0.392)	(0.397)
Median hourly pay (2005)								-1.454**	-1.374**	-1.399**	-1.387**	-1.410**	-1.412**
								(0.665)	(0.661)	(0.658)	(0.657)	(0.683)	(0.683)
Median hourly pay change (2005-2015)						-0.839**	-0.935***	-1.261***	-1.314***	-1.274***	-1.267***	-1.280***	-1.282***
						(0.337)	(0.339)	(0.372)	(0.367)	(0.370)	(0.371)	(0.362)	(0.369)
Unemployment rate (2015)					0.881***	0.897***	0.996***	0.732**	0.842***	0.820***	0.838***	0.862**	0.860**
					(0.299)	(0.301)	(0.304)	(0.304)	(0.305)	(0.307)	(0.314)	(0.336)	(0.344)
Self-employment rate (2015)													-0.016
													(0.380)
Participation rate (2015)												0.069	0.071
												(0.393)	(0.388)
Best subset									X				
Observations	380	380	380	380	377	377	377	377	377	377	377	377	377
R^2	.454	.554	.637	.655	.66	.666	.67	.672	.677	.678	.678	.678	.678

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 6: Predictors of Brexit Vote - Blocked Variable Selection Approach

	Benchmark			Different Best Subsets				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EP election UKIP vote share (2014)	5.954***		5.843***	7.514***				
	(0.292)		(0.352)	(0.294)				
EP election British National Party vote share (2014)	2.071***		2.075***	2.544***				
	(0.212)		(0.232)	(0.257)				
EP election Conservative Party vote share (2014)	1.575***		1.506***	-1.694***				
	(0.356)		(0.392)	(0.341)				
EP election Labour Party vote share (2014)	0.572		0.828*	-1.413***				
	(0.452)		(0.444)	(0.502)				
EP election Green Party vote share (2014)	-1.379***		-1.356***	-2.283***				
	(0.197)		(0.238)	(0.258)				
EP election Lib-Dem vote share (2014)	0.389*		0.417*	-0.369				
	(0.217)		(0.231)	(0.261)				
Initial EU accession migrant share (2001)	-0.697**	-1.724***	-0.714**		-1.793***			
	(0.316)	(0.595)	(0.306)		(0.637)			
EU accession migrant growth (2001-2011)	0.476**		0.481*		1.623***			
	(0.223)		(0.258)		(0.570)			
Initial EU 15 migrant share (2001)		2.751***	0.469		-2.945**			
		(0.453)	(0.346)		(1.164)			
EU 15 migrant growth (2001-2011)			-0.232		-1.609**			
			(0.383)		(0.676)			
Total economy EU dependence		1.267***	0.013		3.460***			
		(0.219)	(0.207)		(0.405)			
Share commuting to London (2011)		1.004**	0.445			-2.695***		
		(0.427)	(0.385)			(0.549)		
Owned (outright + mortgage) share (2001)		2.720***	0.566			6.120***		
		(0.465)	(0.391)			(0.863)		
Council rented share (2001)	0.311		0.554**			1.609***		
	(0.207)		(0.229)			(0.609)		
Total fiscal cuts		-1.626***	-0.368			5.619***		
		(0.444)	(0.323)			(0.488)		
Share suspected cancer patients treated within 62 days (2015)			0.008			-2.398***		
			(0.137)			(0.510)		
Public employment share (2009)		-0.457*	-0.068			-2.278***		
		(0.251)	(0.169)			(0.583)		
Share of pop. no Qualification (2001)	1.056**	6.860***	1.092*				6.649***	
	(0.463)	(0.666)	(0.605)				(0.884)	
Share of pop. no qualification growth (2001-2011)	1.044**	2.577***	0.952*				5.355***	
	(0.451)	(0.453)	(0.503)				(0.582)	
Share of pop. qualification 4+ (2001)	-2.361***	-5.715***	-2.551***				-6.106***	
	(0.409)	(0.727)	(0.544)				(0.827)	
Share of pop. qualification 4+ growth (2001-2011)	-0.502**		-0.506**				2.139***	
	(0.232)		(0.256)				(0.473)	
Share of pop. 60 older growth (2001-2011)			0.120				2.466***	
			(0.197)				(0.290)	
Retail employment share (2001)		1.117***	-0.142					3.938***
		(0.322)	(0.220)					(0.414)
Retail employment share change (2001-2011)		-0.754**	-0.042					-1.389***
		(0.292)	(0.192)					(0.385)
Manufacturing employment share (2001)			-0.108					4.012***
			(0.202)					(0.403)
Manufacturing employment share change (2001-2011)	0.188		0.149					0.818**
	(0.122)		(0.134)					(0.392)
Construction employment share (2001)	0.791***		0.734***					2.977***
	(0.207)		(0.237)					(0.457)
Finance employment share (2001)	-0.603***	-0.850***	-0.546**					1.088**
	(0.179)	(0.315)	(0.228)					(0.464)
Median hourly pay (2005)		1.540***	-0.685*					-1.374**
		(0.407)	(0.350)					(0.661)
Median hourly pay change (2005-2015)			-0.276**					-1.314***
			(0.132)					(0.367)
Unemployment rate (2015)	0.211	0.516**	0.281*					0.842***
	(0.138)	(0.260)	(0.146)					(0.305)
Observations	377	375	374	380	380	375	380	377
R^2	.959	.876	.96	.919	.457	.544	.771	.677

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 5 groups of variables analyzed in Tables 1 through 5. Column 2 shows best subset across all groups of variables except those used in Table 1. Column 3 is full specification based on best subsets determined in Tables 1 through 5. For comparison, columns 4 through 8 re-display the optimal specifications from Tables 1 through 5. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: 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	4.505***					
	(1.343)					
Income Deprivation: Average Rank		5.409***				
		(1.348)				
Employment Deprivation: Average Rank			6.057***			
			(1.264)			
Education and Skills Deprivation: Average Rank				8.259***		
				(1.168)		
Health Deprivation: Average Rank					5.674***	
					(1.326)	
Crime Severity: Average Rank						1.805
						(1.255)
Best subset						
Observations	107	107	107	107	107	107
R^2	.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. The columns present empirical models selected using best subset selection on the set of groups 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 8: 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
R^2	.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 effect exercises, interacting pre-determined "stock" variables measured in 2001 (for the share of households with no qualification, the share of employment in manufacturing and finance) and in 2005 for the median wage with a range of "flow" variables capturing migration 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 9: Did Bad Weather Affect the Referendum Result?

	Turnout		Pct Leave	
	Rainfall		Rainfall	
	Total Amount	Dummy Top Decile	Total Amount	Dummy 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
R^2	.137	.070	.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 *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A Appendix Figures and Tables

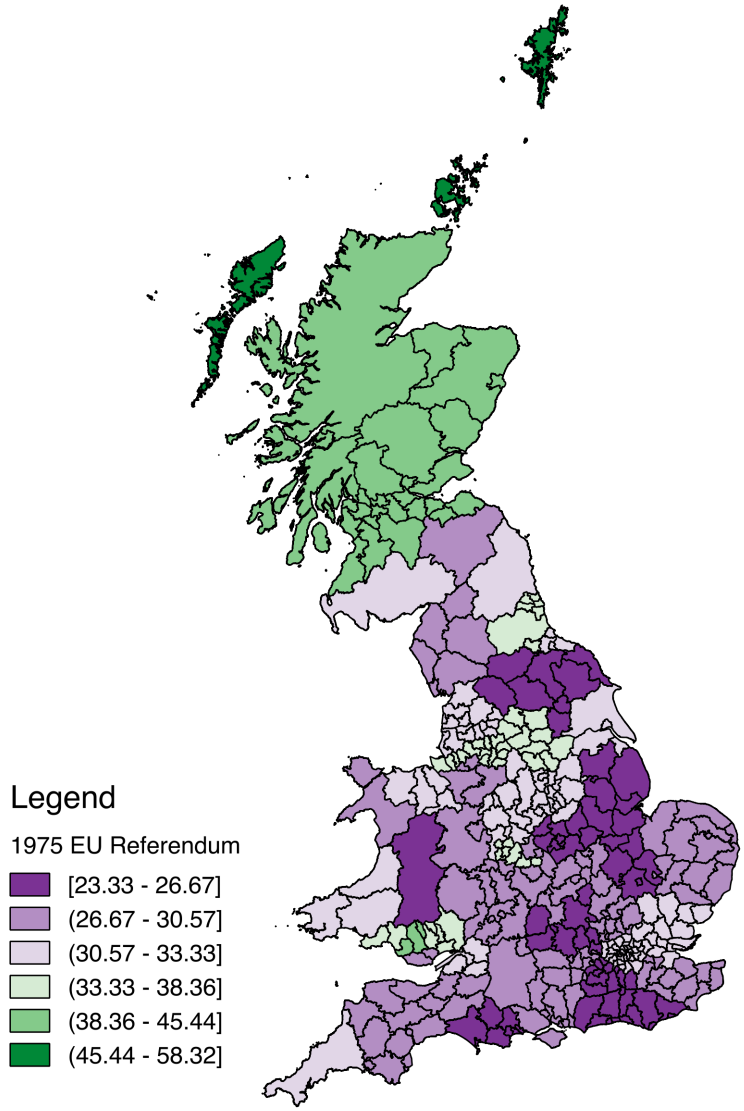

Figure A1: Map of the Leave vote in the 1975 EU Referendum.

Figure A2: Location of cities used for the ward level analysis of Leave support in the EU referendum.

Table A1: Predictors of Referendum Turnout: 1975 Referendum and 2014 European Parliament Elections

	Benchmark							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EP election UKIP vote share (2014)				0.971***	1.164***	1.334***	1.340***	1.380***
				(0.188)	(0.200)	(0.222)	(0.221)	(0.240)
EP election British National Party vote share (2014)						-0.334	-0.300	-0.304
						(0.209)	(0.230)	(0.230)
EP election Conservative Party vote share (2014)	4.076***	2.365***	1.964***	1.833***	2.004***	2.007***	2.011***	2.055***
	(0.179)	(0.304)	(0.291)	(0.262)	(0.286)	(0.290)	(0.291)	(0.343)
EP election Labour Party vote share (2014)		-2.187***	-2.296***	-1.886***	-1.515***	-1.289***	-1.298***	-1.266***
		(0.279)	(0.276)	(0.248)	(0.308)	(0.352)	(0.355)	(0.375)
EP election Green Party vote share (2014)							0.091	0.108
							(0.166)	(0.174)
EP election Lib-Dem vote share (2014)					0.434***	0.458***	0.431**	0.435**
					(0.166)	(0.167)	(0.173)	(0.176)
EP election turnout (2014)			0.843***	1.196***	1.112***	1.039***	1.024***	1.027***
			(0.151)	(0.175)	(0.174)	(0.180)	(0.187)	(0.186)
1975 Referendum Leave share								0.064
								(0.196)
Best subset						X		
Observations	380	380	380	380	380	380	380	380
R^2	.65	.722	.746	.772	.776	.778	.778	.778

Notes: Table reports results from OLS regressions. The columns present empirical models selected using best subset selection on the set of groups 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 A2: Predictors of Referendum Turnout: EU Exposure (Immigration, Trade and Structural Funds)

	Benchmark							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Initial EU accession migrant share (2001)							0.380	0.411
							(0.283)	(0.291)
EU accession migrant growth (2001-2011)								-0.324
								(0.278)
Initial EU 15 migrant share (2001)			0.954*	2.130***	2.179***	2.364***	2.238***	2.177***
			(0.564)	(0.717)	(0.719)	(0.709)	(0.733)	(0.733)
EU 15 migrant growth (2001-2011)				-1.999***	-1.779***	-1.540***	-1.496***	-1.517***
				(0.395)	(0.392)	(0.421)	(0.435)	(0.426)
Initial non-EU migrant share (2001)						-0.769**	-1.001**	-0.828
						(0.376)	(0.434)	(0.510)
Non-EU migrant growth (2001-2011)	-2.014***	-2.248***	-2.663***	-1.986***	-1.881***	-1.534***	-1.529***	-1.433***
	(0.206)	(0.201)	(0.307)	(0.302)	(0.312)	(0.371)	(0.371)	(0.357)
Total economy EU dependence					0.757***	0.708**	0.733***	0.757***
					(0.277)	(0.280)	(0.282)	(0.279)
EU Structural Funds per capita (2013)		-1.685***	-1.546***	-1.397***	-1.044***	-1.091***	-1.076***	-1.080***
		(0.283)	(0.275)	(0.269)	(0.292)	(0.299)	(0.297)	(0.299)
Best subset						X		
Observations	380	369	369	369	369	369	369	369
R^2	.159	.263	.291	.351	.365	.37	.372	.374

Notes: Table reports results from OLS regressions. The columns present empirical models selected using best subset selection on the set of groups 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: Predictors of Referendum Turnout: Public Service Provision and Fiscal Consolidation

	Benchmark							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share commuting to London (2011)						0.406**	0.407**	0.389*
						(0.202)	(0.199)	(0.200)
Owned (outright + mortgage) share (2001)		2.058***	1.253***	1.280***	1.130***	1.455***	1.459***	1.434***
		(0.207)	(0.248)	(0.230)	(0.226)	(0.288)	(0.284)	(0.291)
Owned (outright + mortgage) share growth (2001-2011)					0.552***	0.625***	0.664***	0.685***
					(0.169)	(0.180)	(0.190)	(0.196)
Council rented share (2001)			-1.214***	-1.054***	-1.352***	-1.248***	-1.153***	-1.174***
			(0.223)	(0.211)	(0.218)	(0.230)	(0.260)	(0.268)
Council rented share growth (2001-2011)							0.195	0.179
							(0.201)	(0.202)
Total fiscal cuts	-3.871***	-2.957***	-2.744***	-2.484***	-2.345***	-2.341***	-2.321***	-2.313***
	(0.188)	(0.163)	(0.161)	(0.172)	(0.170)	(0.178)	(0.177)	(0.176)
Share suspected cancer patients treated within 62 days (2015)								-0.140
								(0.174)
Public employment share (2009)				-0.723***	-0.825***	-0.782***	-0.773***	-0.763***
				(0.171)	(0.167)	(0.178)	(0.181)	(0.194)
Best subset						X		
Observations	379	379	379	379	379	376	376	375
R^2	.586	.716	.741	.756	.766	.77	.771	.772

Notes: Table reports results from OLS regressions. The columns present empirical models selected using best subset selection on the set of groups 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: Predictors of Referendum Turnout: Demography and Education

	Benchmark					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of pop. no Qualification (2001)					-0.415	-0.554
					(0.340)	(0.440)
Share of pop. no qualification growth (2001-2011)			1.471***	1.725***	1.443***	1.452***
			(0.172)	(0.169)	(0.299)	(0.300)
Share of pop. qualification 4+ (2001)						-0.163
						(0.326)
Share of pop. qualification 4+ growth (2001-2011)		2.300***	2.350***	2.334***	2.167***	2.152***
		(0.228)	(0.188)	(0.176)	(0.252)	(0.256)
Share of pop. 60 older (2011)				1.160***	1.231***	1.223***
				(0.189)	(0.207)	(0.205)
Share of pop. 60 older growth (2001-2011)	3.385***	2.628***	2.820***	2.060***	1.999***	1.978***
	(0.189)	(0.237)	(0.209)	(0.194)	(0.195)	(0.199)
Best subset					X	
Observations	380	380	380	380	380	380
R^2	.448	.633	.716	.741	.743	.743

Notes: Table reports results from OLS regressions. The columns present empirical models selected using best subset selection on the set of groups 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: Predictors of Referendum Turnout: Economic Structure, Wages and Unemployment

	Benchmark												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Retail employment share (2001)							0.808***	0.678***	0.626***	0.669***	0.608***	0.591***	0.572**
							(0.218)	(0.212)	(0.220)	(0.224)	(0.227)	(0.225)	(0.229)
Retail employment share change (2001-2011)		-1.980***	-1.780***	-1.548***							-0.422**	-0.389*	-0.386*
		(0.219)	(0.201)	(0.213)							(0.202)	(0.210)	(0.210)
Manufacturing employment share (2001)								0.528***	0.613***	0.504**	0.481**	0.612***	0.603***
								(0.201)	(0.203)	(0.209)	(0.209)	(0.231)	(0.230)
Manufacturing employment share change (2001-2011)												0.339	0.338
												(0.234)	(0.234)
Construction employment share (2001)									0.528**	0.593***	0.568***	0.535**	0.543**
									(0.210)	(0.214)	(0.215)	(0.213)	(0.213)
Construction employment share change (2001-2011)					-1.352***	-1.231***	-1.316***	-1.260***	-1.218***	-1.183***	-1.004***	-0.833***	-0.840***
					(0.185)	(0.186)	(0.181)	(0.179)	(0.181)	(0.181)	(0.186)	(0.227)	(0.229)
Finance employment share (2001)										-0.459*	-0.466**	-0.422*	-0.411*
										(0.236)	(0.236)	(0.237)	(0.239)
Finance employment share change (2001-2011)					-1.539***	-1.363***	-1.121***	-1.126***	-0.928***	-0.862***	-0.707***	-0.739***	-0.734***
					(0.210)	(0.225)	(0.218)	(0.223)	(0.228)	(0.232)	(0.237)	(0.236)	(0.237)
Median hourly pay (2005)					1.526***	1.307***	1.610***	1.765***	1.945***	2.217***	2.074***	2.092***	2.043***
					(0.232)	(0.237)	(0.253)	(0.270)	(0.279)	(0.278)	(0.288)	(0.292)	(0.298)
Median hourly pay change (2005-2015)													-0.064
													(0.232)
Unemployment rate (2015)	-2.721***	-2.216***	-1.888***	-1.412***		-1.060***	-0.997***	-1.008***	-1.083***	-1.068***	-1.085***	-1.071***	-1.071***
	(0.242)	(0.224)	(0.214)	(0.211)		(0.207)	(0.200)	(0.198)	(0.197)	(0.195)	(0.195)	(0.195)	(0.195)
Self-employment rate (2015)			1.412***	1.241***	1.486***	1.402***	1.466***	1.607***	1.610***	1.588***	1.518***	1.491***	1.484***
			(0.203)	(0.199)	(0.190)	(0.192)	(0.190)	(0.201)	(0.197)	(0.197)	(0.195)	(0.194)	(0.194)
Participation rate (2015)				1.269***	1.537***	1.172***	1.202***	1.200***	1.131***	1.110***	1.068***	1.045***	1.057***
				(0.241)	(0.238)	(0.234)	(0.231)	(0.230)	(0.227)	(0.225)	(0.229)	(0.229)	(0.231)
Best subset												X	
Observations	377	377	377	377	378	377	377	377	377	377	377	377	377
R^2	.289	.433	.498	.543	.56	.592	.609	.616	.623	.627	.63	.633	.633

Notes: Table reports results from OLS regressions. The columns present empirical models selected using best subset selection on the set of groups 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

	Benchmark			Different Best Subsets				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EP election UKIP vote share (2014)	1.013***		1.868***	1.334***				
	(0.230)		(0.261)	(0.222)				
EP election British National Party vote share (2014)	-0.629***		-0.523***	-0.334				
	(0.154)		(0.185)	(0.209)				
EP election Conservative Party vote share (2014)	0.661***		0.906***	2.007***				
	(0.246)		(0.305)	(0.290)				
EP election Labour Party vote share (2014)			1.012***	-1.289***				
			(0.361)	(0.352)				
EP election Lib-Dem vote share (2014)			0.392***	0.458***				
			(0.149)	(0.167)				
EP election turnout (2014)			1.201***	1.039***				
			(0.149)	(0.180)				
Initial EU 15 migrant share (2001)	0.072	0.590	0.054		2.364***			
	(0.389)	(0.385)	(0.245)		(0.709)			
EU 15 migrant growth (2001-2011)	-0.049		0.070		-1.540***			
	(0.243)		(0.319)		(0.421)			
Initial non-EU migrant share (2001)			-0.518		-0.769**			
			(0.384)		(0.376)			
Non-EU migrant growth (2001-2011)	-0.936***	-1.094***	-0.834***		-1.534***			
	(0.236)	(0.251)	(0.260)		(0.371)			
Total economy EU dependence		-0.169	-0.369***		0.708**			
		(0.130)	(0.132)		(0.280)			
EU Structural Funds per capita (2013)		-0.169	0.106		-1.091***			
		(0.151)	(0.139)		(0.299)			
Share commuting to London (2011)			0.103			0.406**		
			(0.314)			(0.202)		
Owned (outright + mortgage) share (2001)	0.597*		0.417			1.455***		
	(0.319)		(0.276)			(0.288)		
Owned (outright + mortgage) share growth (2001-2011)		-0.664***	-0.111			0.625***		
		(0.165)	(0.205)			(0.180)		
Council rented share (2001)	-0.098		-0.075			-1.248***		
	(0.194)		(0.163)			(0.230)		
Total fiscal cuts	-1.286***	-1.797***	-1.404***			-2.341***		
	(0.267)	(0.272)	(0.243)			(0.178)		
Public employment share (2009)		-0.390**	-0.312**			-0.782***		
		(0.151)	(0.126)			(0.178)		
Share of pop. no Qualification (2001)	0.401	0.521	-0.773				-0.415	
	(0.499)	(0.322)	(0.476)				(0.340)	
Share of pop. no qualification growth (2001-2011)	0.390		-0.560*				1.443***	
	(0.362)		(0.326)				(0.299)	
Share of pop. qualification 4+ growth (2001-2011)	1.113***	1.166***	0.289				2.167***	
	(0.195)	(0.164)	(0.188)				(0.252)	
Share of pop. 60 older (2011)		1.129***	0.450**				1.231***	
		(0.196)	(0.220)				(0.207)	
Share of pop. 60 older growth (2001-2011)	0.561***	0.757***	0.146				1.999***	
	(0.180)	(0.204)	(0.179)				(0.195)	
Retail employment share (2001)		0.542***	0.178					0.591***
		(0.169)	(0.144)					(0.225)
Retail employment share change (2001-2011)			0.101					-0.389*
			(0.177)					(0.210)
Manufacturing employment share (2001)			0.328*					0.612***
			(0.176)					(0.231)
Manufacturing employment share change (2001-2011)	0.276***	0.312***	0.372***					0.339
	(0.095)	(0.097)	(0.139)					(0.234)
Construction employment share (2001)	-0.458***	-0.417***	-0.301**					0.535**
	(0.146)	(0.156)	(0.150)					(0.213)
Construction employment share change (2001-2011)			0.196					-0.833***
			(0.172)					(0.227)
Finance employment share (2001)			-0.389***					-0.422*
			(0.134)					(0.237)
Finance employment share change (2001-2011)		-0.452**	-0.280					-0.739***
		(0.180)	(0.188)					(0.236)
Median hourly pay (2005)	0.517**	1.058***	0.741***					2.092***
	(0.202)	(0.202)	(0.199)					(0.292)
Unemployment rate (2015)			-0.114					-1.071***
			(0.098)					(0.195)
Self-employment rate (2015)		0.268**	0.231**					1.491***
		(0.123)	(0.111)					(0.194)
Participation rate (2015)	0.257*	0.274*	0.157					1.045***
	(0.152)	(0.140)	(0.126)					(0.229)
Observations	379	368	366	380	369	376	380	377
R^2	.862	.871	.914	.778	.37	.77	.743	.633

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 5 groups of variables analyzed in Tables A1 through A5. Column 2 shows best subset across all groups of variables except those used in Table 1. Column 3 is full specification based on best subsets determined in Tables A1 through A5. For comparison, columns 4 through 8 re-display the optimal specifications from Tables A1 through A5. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A7: Exploiting Cross-Sectional Variation in Index of Deprivation Ranks Across 107 Wards in 4 Cities

	(1) Leave	(2) Leave	(3) Leave	(4) Leave	(5) Leave	(6) Leave	(7) Leave	(8) Leave	(9) Leave	(10) Leave	(11) Leave	(12) Leave
Index of Multiple Deprivation: Average Rank					-8.581***						2.203	6.025
					(2.192)						(7.160)	(9.376)
Income Deprivation: Average Rank						-7.520***	-23.879***	-7.054***	-8.648***	-15.504***	-16.681***	-17.245***
						(2.106)	(5.428)	(1.780)	(2.149)	(5.427)	(6.212)	(6.033)
Employment Deprivation: Average Rank							15.418***			6.535	6.460	5.886
							(4.537)			(4.819)	(4.934)	(5.164)
Education and Skills Deprivation: Average Rank	8.494***	8.309***	7.890***	7.567***	14.941***	14.057***	15.118***	16.169***	16.330***	16.465***	15.979***	15.645***
	(1.139)	(1.109)	(0.795)	(0.776)	(1.971)	(1.970)	(1.873)	(1.895)	(1.976)	(1.945)	(2.312)	(2.482)
Health Deprivation: Average Rank												-2.056
												(3.569)
Crime Severity: Average Rank									1.894	1.722	1.387	1.146
									(1.537)	(1.540)	(1.868)	(2.055)
Barriers to Housing and Services: Average Rank								-5.517***	-5.586***	-4.795**	-5.131**	-5.618**
								(1.617)	(1.677)	(1.840)	(2.483)	(2.430)
Living Environment Deprivation: Average Rank			-7.916***	-7.841***	-5.773***	-7.078***	-5.840***	-6.045***	-6.599***	-6.171***	-6.415***	-6.871***
			(0.819)	(0.737)	(0.977)	(0.808)	(0.896)	(0.804)	(0.907)	(0.983)	(1.246)	(1.453)
city==Birmingham		4.888*	10.950***	8.110***	8.097***	6.509***	1.331	26.259***	28.349***	27.204***	27.366***	16.318***
		(2.579)	(1.914)	(2.041)	(2.007)	(1.990)	(2.621)	(3.187)	(3.650)	(3.780)	(4.048)	(3.760)
city==Bristol				-5.929***	-9.367***	-11.690***	-17.316***	0.000	0.000	0.000	0.000	-11.452***
				(1.478)	(1.882)	(2.516)	(3.083)	(.)	(.)	(.)	(.)	(2.182)
city==Greenwich								18.000***	17.836***	19.339***	19.810***	7.863*
								(3.088)	(3.262)	(3.217)	(3.876)	(4.637)
city==Nottingham						-3.413	-8.663***	10.497***	11.138***	10.924***	10.791***	0.000
						(2.298)	(3.042)	(2.054)	(2.092)	(2.111)	(2.116)	(.)
Best subset									X			
Observations	107	107	107	107	107	107	107	107	107	107	107	107
R^2	.352	.377	.638	.666	.705	.707	.728	.745	.75	.753	.754	.755

Notes: Table reports results from OLS regressions. The dependent variable is the Vote Leave share in one of the 107 wards of the four cities Birmingham, Bristol, Nottingham and Greenwich/London. 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 A8: Exploiting Cross-Sectional Variation in Index of Deprivation Ranks Across 107 Wards in 4 Cities

	(1) Leave	(2) Leave	(3) Leave	(4) Leave	(5) Leave	(6) Leave	(7) Leave
Income Deprivation: Average Rank						-10.078*	-8.651
						(5.847)	(5.835)
Employment Deprivation: Average Rank					8.318***	14.674***	14.153***
					(3.163)	(4.347)	(4.435)
Education and Skills Deprivation: Average Rank	8.494***	8.349***	9.915***	8.800***	12.895***	11.766***	11.733***
	(1.139)	(0.886)	(1.195)	(1.180)	(2.362)	(2.596)	(2.614)
Health Deprivation: Average Rank							1.494
							(2.606)
Crime Severity: Average Rank			-2.763**	-2.673**			
			(1.247)	(1.254)			
Barriers to Housing and Services: Average Rank				2.252**	4.554***	5.011***	5.165***
				(0.992)	(1.314)	(1.303)	(1.342)
Living Environment Deprivation: Average Rank		-6.030***	-5.453***	-6.226***	-3.993**	-4.662***	-4.456***
		(0.995)	(1.069)	(1.058)	(1.568)	(1.664)	(1.647)
Best subset						X	
Observations	107	107	107	107	107	107	107
R^2	.352	.526	.549	.565	.58	.588	.589

Notes: Table reports results from OLS regressions. The dependent variable is the Vote Leave share at the ward level across four English cities. The columns present empirical models selected using best subset selection on the set of groups predictors using the AIC information criterion. Robust standard errors are presented in parentheses, asterisks indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.