Multilingualism and the Brexit Referendum

Roland Kappe

Department of Political Science

University College London

`r.kappe@ucl.ac.uk`

This version: January 2020[*]

Abstract

This paper argues that the (lack of) foreign language skills has contributed to the outcome of the Brexit referendum. Theory suggests that speaking foreign languages reduces perceptions of cultural distance and contributes to the formation of transnational identities. Research also shows a link between language skills and European identity (Kuhn, 2015; Díez Medrano, 2018). Did Britons' relative lack of foreign language skills play a role in the Brexit decision? The data show that those with foreign language skills overwhelmingly voted to remain. But is this a genuine effect, or are foreign languages skills simply a marker of e.g. higher educational attainment which has been found to predict the individual Brexit vote? Using matching methods and data from the referendum wave of the British Election Study, it is possible to estimate the effect of foreign language skills on the referendum vote. The results suggest that a significant effect of foreign language skills remains, even when taking into account various factors such as education, age, gender, income, and region, as well as party preference and personality differences.

Keywords: Foreign language skills, European identity, Brexit
JEL Codes: D72, D91, Z13

1 Introduction

This paper provides a case study of the role of multilingualism in politics by looking at the relationship between multilingualism and Brexit. Political science – and specifically the study of public opinion and political behaviour – often use data from nationally representative surveys to make inferences about what explains people's political behaviour and attitudes. In this case, data from the referendum wave of the British Election Study allow us to investigate the relationship between multilingualism and the vote in the Brexit referendum.

Compared to citizens of other European countries, fewer Britons speak foreign languages. More than 65% of Britons aged 25–64 know no foreign languages at all. Figure 1 shows the data for the UK and other EU member states (Eurostat, 2016).
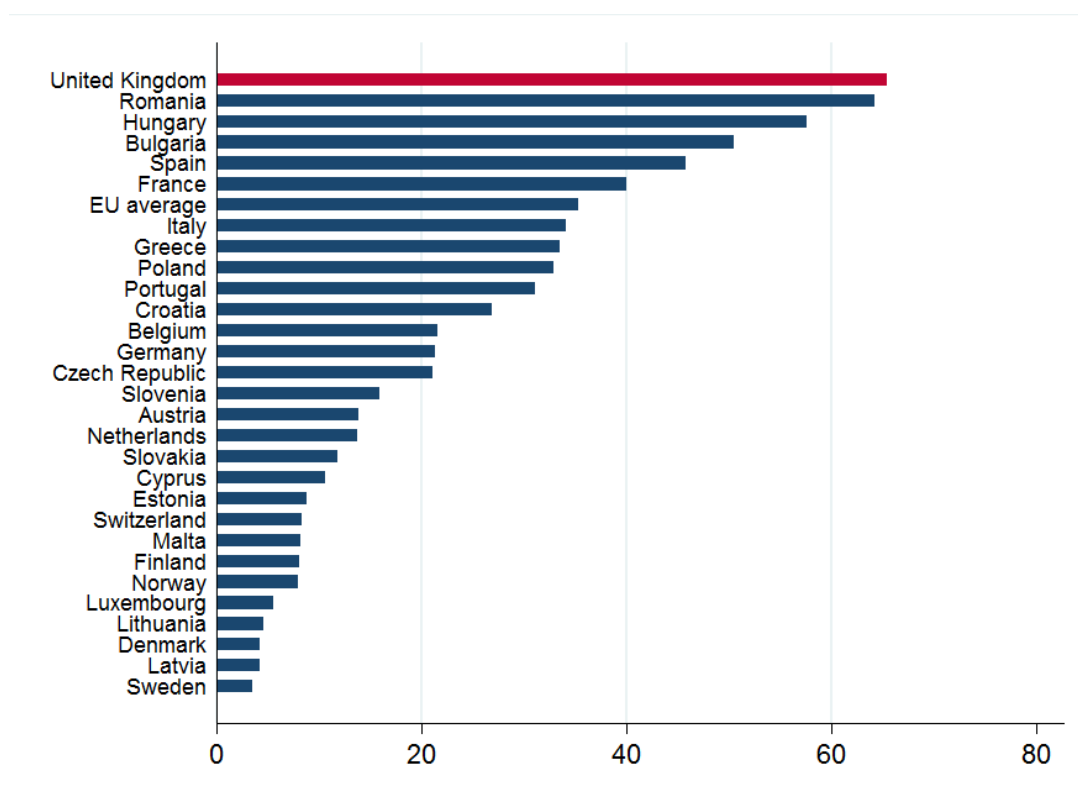

Figure 1: Percentage of people aged 25-64 who speak no foreign language at all

Is it possible that this relative lack of foreign language skills contributed to the outcome of the Brexit referendum? Theoretical work suggests that speaking foreign languages reduces perceptions of cultural distance and contributes to the formation of

transnational identities (Benet-Martínez and Haritatos, 2005; Kuhn, 2011). Recent research also shows a link between language skills and European identity (Kuhn, 2015; Díez Medrano, 2018). Applying this theoretical framework to the case of the "Brexit referendum" in the UK leads to the hypothesis that there is a relationship between language skills and the Brexit vote.

This paper uses data from the referendum wave of the British Election Study (Fieldhouse et al., 2017) to test this hypothesis empirically. The data show that in the June 2016 referendum on membership in the European Union, those with foreign language skills voted in favour to remain a member of the EU (58 %), while those who speak no second language voted 54 to 46 per cent in favour of Brexit. This leads to the question whether this is a genuine effect, or whether foreign languages skills are essentially epiphenomenal, and simply a marker of education more generally, which is a well-established contributor to the Brexit decision (Goodwin and Heath, 2016; Clarke et al., 2017). To answer this question, the paper uses matching methods to take into account factors such as education, age, income and location, and estimate the effect of foreign language skills on the Brexit decision. The results show that even after matching on a wide range of plausible confounders, language skills have an effect on individuals' decision to vote for or against Brexit.

The paper hopes to contribute to the literature on language and politics, specifically multilingualism and political identity, as well as the work on European identity, Euroscepticism and the causes of Brexit. Firstly, because the Brexit referendum provides an important test case, as the vote decision is consequential, going beyond just responses to a survey-based attitudinal measure of European identification. Secondly, while a variety of demographic, personality and attitudinal factors have been linked to the Brexit decision (Clarke et al., 2017), multilingualism (or the lack of it) has so far not been discussed as a contributing factor. Furthermore, questions of language go beyond the established narratives of age, education and the economically or culturally 'left behind' (Goodwin and Heath, 2016). This is especially important as many established factors that contributed to Brexit – demographics and location for example – are difficult or

impossible to influence. Language skills and language learning in contrast are one of the variables that can actually be affected by government policy, e.g. through curriculum reforms or funding choices. This leads to the third contribution which is the recognition that EU policies in support of language learning and exchange are important tools for creating a European identity and cohesive union. The evidence showing that language skills have affected the Brexit referendum provides support for the idea that these programmes are indeed vital.

2   Theoretical Background

Why should there be a relationship between people's language skills and the decision to vote to leave the European Union in the Brexit referendum? This section presents several complementary theoretical perspectives that imply such a link. First, it broadens the lens and discusses some of the wider literature on language and identity that are relevant as a theoretical background. In a second step, it reviews some of the evidence that suggests a direct relationship between language learning and perceptions of cultural distance which in turn could affect how close or distant UK voters feel to other Europeans either through the learning process itself or opportunities for contact. The third part concerns the link between foreign language skills and the possibility of a European public sphere – participation in which may enable the perception of the EU as a legitimate political entity. The final, and arguably most important theoretical mechanism can be found in the literature on European identity, where foreign language skills have been considered as indicative of cosmopolitanism or transnationalism – which in turn contributes to an individual's sense of a European identity. The section concludes with a brief overview of other factors that have been found to be related to the leave vote and that could potentially confound the effect of foreign language skills on the Brexit decision.

There is a longstanding historical view that a common language and national identity are intricately linked (Anderson, 1983; Hobsbawm, 1990). A common language facilitates the formation of a shared culture through discourse and the construction of cultural mean-

ing that is shared by members of a community. Shared cultural constructs in turn can form the core of a national identity. As proto-national communities exceed the experiential sizes of groups in traditional societies, the construction of shared culture happens through for example the printing press and a mediated public sphere. Anderson (1983) describes this process of emergence, and labels nation states as "imagined communities" because their members cannot all interact directly. Using the Austrian case, Wodak et al. (2009) nicely unpack the micro-processes that underlie this construction of national identity. At the same time, common language and culture then also allow for the formation of a perceived in-group that becomes visible mainly in contrast to out-groups that do not share either the cultural constructs, or the language, or both. This links the literature on nationalism and identity to the social psychology literature discussed below. The relationship is not one-directional, however: In the nineteenth and twentieth century, nation states also actively promoted language homogeneity (Blommaert and Verschueren, 1992; Wright, 2016). So language can be seen as both a part of the origin and as a consequence of the formation of the modern nation state.

The centrality of language to national identity persists today. Countries regularly require speaking the national language as a condition for residence or citizenship, and a recent analysis of cross-national data from Pew Research Center's "Global Attitudes Survey" shows that in most countries, people say that speaking the language is more important to national identity than someone's birthplace (Stokes, 2017). It is perhaps not surprising then, that following a more open, transnational period, a nationalist backlash finds fertile ground in monolingual environments. This is, in part, what this chapter attempts to show using the Brexit case study.

Social psychologists, linguists, and education researchers have also analysed the relationship between language and identity. A large body of work explores these questions particularly as they relate to multilingual societies (Pavlenko and Blackledge, 2004), or immigration experiences, and the creation of amalgamated identities, e.g. Latinos in the US case (e.g. Padilla and Perez, 2003; Schecter and Bayley, 2005). One of the major contributions from cross-cultural psychology in this respect is bicultural identity

integration theory (Benet-Martínez et al., 2002; Benet-Martínez and Haritatos, 2005). Benet-Martínez and co-authors unpack the process underlying the integration of dual cultural identities, and analyse how these processes relate to perceptions of identity and other cognitive processes such as 'cultural frame switching' and 'code switching'. For the question at hand, effects of foreign language learning outside an immigration or minority context e.g. in school are perhaps more relevant, and the relationship between language instruction and learning about the culture of speakers of the language have been explored in education research (Byram, 1994). The fundamental idea is that learning a foreign language usually also exposes the student to the culture of the people whose language is being studied[1].

Finally, speaking a foreign language also increases opportunities for direct interaction with speakers of this language, either with immigrants in one's own country, or abroad, for example on holiday. There is strong evidence that direct contact with members of an out-group reduces prejudice and negative attitudes towards this group (Allport, 1954). This intergroup contact hypothesis has been expanded and replicated in various settings (see Pettigrew and Tropp (2006) for a meta-analysis). Interestingly, there is even some evidence that simply imagining interactions with a member of another group can lead to more positive perceptions of this group (Crisp and Turner, 2009). If true, this of course points to the possibility of activities like role plays in a language classroom, and cultural learning more generally, affecting outgroup views. Overall, speaking a foreign language increases opportunities for direct intergroup contact which reduces prejudice, and it has also been shown to reduce perceptions of cultural distance between one's native culture and speakers of the other language (Benet-Martínez and Haritatos, 2005). In turn, this reduction in prejudice and increased perception of proximity towards other Europeans among speakers of foreign languages in the UK could then plausibly have had an effect on the Brexit vote.

Narrowing the focus from more general questions of language and identity to research that directly relates to the particular connection of foreign language skills and

---

[1]For an in-depth theoretical framework and case studies of the role of culture in language learning see Byram (1994).

the European Union, two topics stand out as having received widespread attention in the literature: The question of a common language as a precondition in the development of a European public sphere on the macro level, and the role of language skills in the formation of a European identity on the micro level. These are examined below.
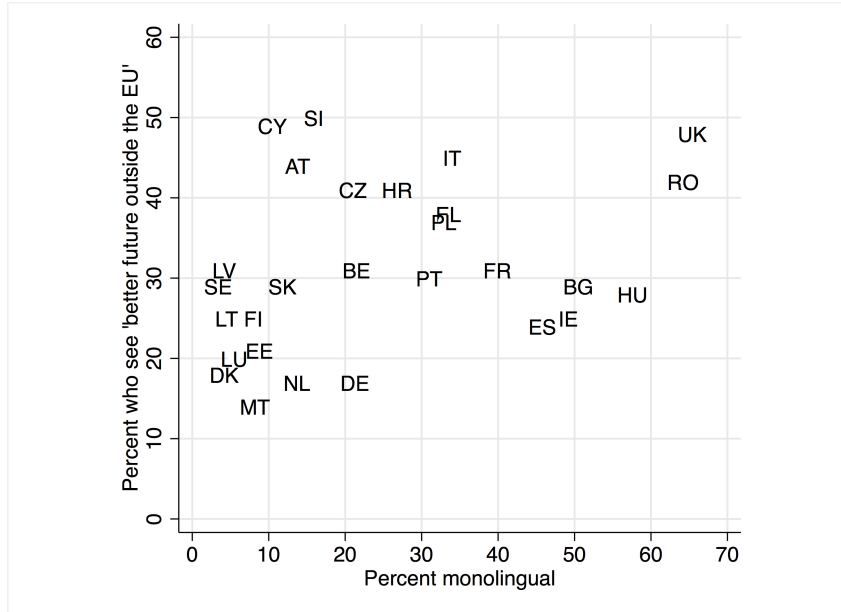
The question of the emergence of a European public sphere is of particular interest for a study on language and Brexit. The basic idea is that for the European Union to be perceived by citizens as a legitimate and democratic supranational authority, a European public sphere, in which Europe-wide political debate and public discourse can take place, is required (Risse, 2015a). Public opinion and media discourse about European issues should be transnational, or at least harmonised between the individual member states, rather than compartmentalised and idiosyncratic. The work of Jürgen Habermas (1974, 1981), specifically his theory of communicative action, is the starting point for attempts at definition, systematic theorising, and empirical analysis of the concept of a public sphere, but his work also blends systematic analysis with a normative component[2]. Detaching from this and focusing on a purely analytic perspective, Gerhards and Neidhardt (1990) provide a thorough conceptual systematisation and analysis of emergence, function, structure, role-differentiation, actor-strategies, and public opinion processes of a 'public sphere'. The concept received much attention by scholars focusing on European Union politics (Eder and Kantner, 2005; Risse, 2003; Trenz and Eder, 2004; Eriksen, 2005; Koopmans, 2007). These scholars at the core try to answer the question of the possibility of a European public sphere, and what constitutes evidence of this emerging public sphere. Thomas Risse's contribution (2015a) on the interplay of the emerging European public sphere and European identity, and a recent edited volume (2015b) offer an excellent overview of the current state of research, while Pana (2015), and Walter (2017) clearly distinguish between different theoretical models of the public sphere that underlie some of these debates. The implicit assumption in much of this field has been that a genuine deliberative European public sphere requires a common language – or common languages – or at least the ability to follow and participate in the discourse in a foreign

---

[2]For important work in this tradition, see for example Calhoun (1996); Fraser (1990); Strani (2010, 2014).

language. This is where the questions surrounding a European public sphere relate to the issue of foreign language skills and Brexit. Being unable to follow the public discourse in say France or Germany limits the opportunity of individual citizens – but in many cases also of elites such as journalists or politicians – to participate in the European public sphere which in turn has negative consequences for views about the EU and support of the European project. Breidbach (2002) discusses the role of foreign language teaching for the development of a European public sphere. For an interesting counterpoint on the assumption of a common language as a necessary condition for a European public sphere, however, see Doerr (2012) who analyses debates and exchange by grassroots activists who intentionally use translation practice as a method of deliberation in the multilingual European Social Forum.

Theories concerning a European public sphere have in common that they are necessarily describing the emergence of a macro-level phenomenon – and a possible feedback mechanism in the form of effects on European identity or Euroscepticism in the aggregate. When looking at average language skills and Euroscepticism on the cross-national aggregate level, however, foreign language skills are also not a panacea. Figure 2 shows the relationship between the percentage of adults aged 25–64 reporting they know no foreign languages (Eurostat, 2016) and the percentage of respondents who see a better future for their country outside the EU (European Commission, 2017). While the UK – with a low level of foreign language skills and high level of support for leaving the EU – fits the expected pattern, there are also several countries with high levels of Euroscepticism despite relatively good average foreign language skills. It is therefore necessary to look beyond the aggregate level, and investigate what this relationship looks like on the individual level.

Do foreign language skills contribute to the formation of a European identity on the individual level? Two potential mechanisms are covered by the literature. The first rests on intergroup contact theory which has already been discussed above (Allport, 1954; Pettigrew and Tropp, 2006). Foreign language skills allow for transnational interactions, and transnational interactions are related to European identity directly (Fligstein, 2008),

7

Notes: Percentage of respondents who agree "[Country] could better face the future outside the EU" from Eurobarometer 86 (European Commission, 2017). Percentage of adults aged 25–64 reporting they know no languages beyond their mother tongue. Data based on the 2016 Adult Education Survey (Eurostat, 2016)

Figure 2: Lack of language skills and Euroscepticism

or lead to generally cosmopolitan attitudes (Mau et al., 2008), which in turn might contribute to an individual's sense of European identity and ultimately support for the European Union (Fligstein, 2008; Kuhn, 2015). The second theoretical mechanism treats foreign language skills as a component of individual transnationalism (Kuhn, 2011), and embeds individual transnationalism in the theoretical framework on security communities by Deutsch et al. (1957). In this view, security communities can be created by increasing and institutionalising cross-border transactions. If they are sustained over time and multidimensional, these institutionalised transactions then increase trust between the countries' populations and support for the security community. Kuhn (2011) applies this logic to the European Union today.

The motivation for Kuhn's (2011) article is the apparent contradiction between increased transnational interactions and European integration on the one hand, and recent increases in Euroscepticism across many member states on the other hand. Following Deutsch et al. (1957), the exact opposite should occur. Transnational interactions should drive support for further integration. Kuhn resolves this puzzle by noticing that these transnational interactions are – for the most part – concentrated in a small part of the

8

population, as also argued by Fligstein (2008). Transnationalism should therefore be measured at the individual level. Kuhn defines three dimensions of individual transnationalism: Direct transnational practices such as stays abroad, having a transnational background, such as being foreign-born, and possessing transnational human capital. While the concept of transnational human capital can be defined widely, foreign language skills are the crucial factor (Kuhn, 2011: 814). Specifically, transnational human capital is operationalised using Eurobarometer survey items on "self-assessed preparedness to buy a product in another EU language", and "having read a newspaper, book, or magazine in a foreign language in the past 12 months" (Kuhn, 2011: 820). Using Eurobarometer data from 2006, her analysis supports the hypothesised relationship between individual transnational human capital (i.e. foreign language skills) and lower Euroscepticism. More recently, Díez Medrano (2018) provides a detailed analysis of the relationship between multilingualism and European identification. He uses Eurobarometer data from 2010 to test the effect of the number of languages an individual speaks on European identity, measured as the answer to "Thinking about the fact that you are European, how important is being European for you personally?" on a four point Likert-scale, and finds a modestly sized but statistically significant effect of about 6% of the outcome variable's inter-quartile-range. In a second step, he uses mediation analysis to see how much of this effect can be attributed to actual interaction (e.g. in the form of holidays, living, or working abroad). The data suggest that while about a third of the effect is attributable to actual transnational interactions, a large part remains unexplained. This opens up the possibility for the more complex socio-psychological effects associated with foreign language learning, multilingualism, and perceptions of cultural distance and transnational identity I presented above.

To summarise, theoretical perspectives and empirical work from diverse disciplines points to a complex relationship between multilingualism and (European) identity. On the macro level historians have explained national identity based on language, and scholars of European politics have discussed the role of a European public sphere. On the micro level, work by social psychologists and linguists points to intergroup contact theory and the

effect of foreign language learning on perceptions of cultural distance, while those working on European political behaviour relate language and European identity via individual transnationalism. All these theoretical mechanisms and the body of evidence supporting them have in common that they lead to an expectation that multilingual individuals exhibit a higher level of European identity. Applying this general relationship to the case of the 2016 Brexit referendum leads to the hypothesis that multilingual UK citizens were more likely to embrace a European identity and should be more likely to have voted to remain in the European Union. The next section outlines how this conjecture can be tested using data from the 2016 referendum wave of the British Election Study, and presents a testable hypothesis.

Various other factors have been put forward to explain the outcome of the 2016 Brexit referendum. While the work discussed above points to the lack of foreign language skills having played a role in the vote, many other factors are likely to have been more influential overall. Furthermore, some of these such as age and education are also related to foreign language skills. It is therefore necessary to take these into account in the following analysis. In terms of factors that have been identified in several analyses of the Brexit referendum, core demographic factors, including education level, income, gender, and age are usually relevant (Hobolt, 2016; Clarke et al., 2017). Factors associated with voters' geographic location, e.g. in debates surrounding areas of 'economic decline' versus areas that have seen immigration and 'cultural backlash', are perhaps the most prominent (Goodwin and Heath, 2016; Carreras et al., 2019). Research also finds effects of party identification due to elite cue taking (Hobolt, 2016; Clarke et al., 2017), as well as – while not directly modelling the Brexit vote, but in terms of support for the Eurosceptic UKIP party – effects of voter personality (Kappe, 2015).

3   Hypothesis, Data, and Methods

The hypothesis is that there is a relationship between language skills and the Brexit vote. The chapter provides an empirical test for this relationship. Using British Election Study

(BES) survey data, it tests whether there is a relationship between self-reported language skills and the Brexit vote.

The referendum wave of the British Election study (Fieldhouse et al., 2017) was conducted between May and June 2016 in the run-up of the Brexit referendum. The survey asked a nationally representative sample of 33,502 respondents about a variety of political attitudes, identifications, their vote intention in the referendum, as well as demographics, including knowledge of other languages. Specifically, it asked respondents whether they "speak a language other than English at conversational level"[3]. The hypothesis thus is:

Hypothesis: *Voters who "speak a language other than English (or Welsh) at conversational level" are more likely to have voted to remain in the European Union in the 2016 referendum.*

The British Election Study only includes this question on language skills. A potential problem with this survey item is misreporting, or an overly generous assessment of what "conversational level" means. While problems with self-reported language skills have been identified in the literature (Edele et al., 2015), no simple alternative is readily available, and misreporting only poses a threat to inference if it is systematically related to the Brexit vote. Furthermore, the question wording is similar to – for example – the Eurobarometer survey which asks respondents about their mother tongue and "what other language(s) do you speak well enough to be able to have a conversation?", which is widely used as a measure of multilingualism (e.g. Díez Medrano, 2018).

As mentioned above, factors other than foreign language skills have already been identified to be associated with vote choice in the Brexit referendum (Hobolt, 2016; Clarke et al., 2017). They include education, age, gender, income, location, as well as family background in terms of minority status or foreign-born parents, personality factors, and party identification. All of these can be measured using variables in the BES dataset and will be accounted for in the following analysis. A complete list of variables and operationalisation can be found in Table A1 in the appendix.

---

[3]For respondents from Wales the question wording is: *"Do you speak a language other than English (or Welsh) at conversational level".*

In terms of summary statistics, the sample corresponds nicely to the referendum outcome with 49% of those intending to vote choosing Remain and 51% Leave. Table 1 in the next section shows the percentage of foreign language speakers and Remain voters by gender, education level, and non-white and parental background in the sample. Table A2 in the appendix provides summary statistics and the correlation with multilingualism and Remain vote for age and personality factors.

The data analysis presented here uses propensity score matching. With matching, it is possible to compare the Brexit vote of people who differ in terms of language skills, but who are otherwise very similar – or identical – in terms of education, income, age, gender, and possibly party identification and personality characteristics. While language skills necessarily pre-date the vote choice, and many plausible confounders can be controlled for using matching, the identification strategy ultimately relies on observational data, which limits the possibility of claiming this to be a well-identified causal effect, as possible unobserved confounders could bias the results. That being said, matching methods allow for a careful controlled comparison, and the analysis tries to cover possible confounders and is clear about the underlying assumptions. The estimated average treatment effects appear statistically significant, robust to the inclusion of additional covariates, and realistic in terms of magnitude.

4 Data Analysis

Did Britons' relative lack of foreign language skills contribute to the outcome of the Brexit referendum? To answer this question, this section tests the hypothesis set out above. Firstly, is there a difference in the referendum vote between those who speak a foreign language and those who only speak English (or Welsh in Wales)? Looking at the data from the referendum wave of the BES, Figure 2 supports the original assumption. Among those who only speak English, 54.6% voted to leave the European Union, while among those who speak an additional language, only 41.3% voted to leave the EU.

This leads to the question whether this is a genuine effect, or whether the factor of
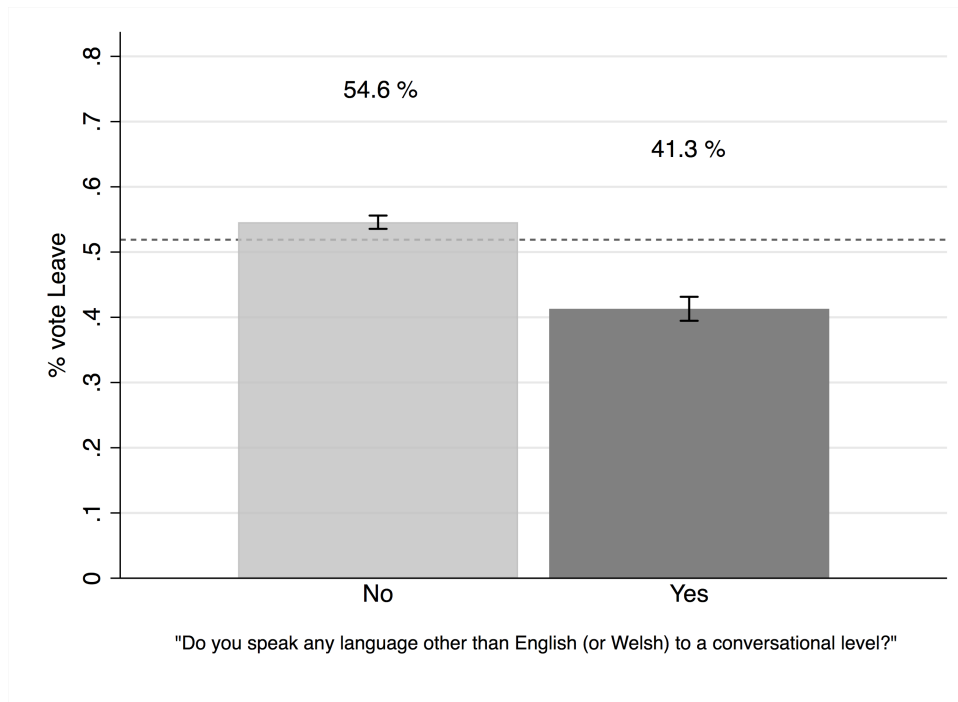
Figure 3: Difference in referendum vote by language skills

foreign languages skills is epiphenomenal. Foreign language skills could – for example – simply be a marker of education more generally which is a well-established contributor to the Brexit decision (Goodwin and Heath, 2016; Clarke et al., 2017). Similarly, those who have foreign-born parents or belong to an ethnic minority, might simultaneously have language skills beyond English, and be less inclined to support Brexit, e.g. due to the nativist and xenophobic messaging of the Leave campaign. In sum, there are a variety of possible confounding variables that could plausibly affect both language skills and the referendum vote.

I am interested in the difference in the referendum vote between those who speak another language and those who do not. The fundamental problem of course is that we cannot observe the outcome for participants both with and without this 'treatment'. Furthermore, the two groups are different in a variety of respects, such as age, education, and other characteristics. Table 1 shows the percentage of foreign language speakers and Remain voters by gender, education level, and ethnic and parental background in the sample.

To address this problem, matching methods can be useful. Matching allows us to

Table 1: Percentage of foreign language speakers and Remain voters by gender, education level, and ethnic and parental background in the BES dataset

	Overall%	Multilingual %	Remain%
Male	49	22	48
Female	51	24	50
Education			
No Qualifications	8.2	6.4	25
GCSE D-G	4.9	8.4	31
GCSE A*-C	22	11	33
A-level	21	21	49
Undergraduate	33	32	62
Postgrad	11	44	73
Non-white background	8.4	56	65
White background	92	19	48
Foreign born parent(s)	13	48	59
UK born parents	87	19	47
All	100	23	49

limit our analysis to a comparison of voters who differ in terms of language skills, but who are otherwise identical – or very similar – in terms of observable confounders such as education, income, age, gender, and possibly other factors such as party identification and personality characteristics.

As an example of how different the two groups are, it is worth looking at education in Table 1. In the sample overall, 23% of UK respondents claim to speak a foreign language at conversational level, but only 6% of those with no qualifications do, while this rises to 44% amongst those with postgraduate educational qualifications. Using matching methods is therefore suited to the data at hand, as in the overall sample, the difference between the treatment and control groups is large in terms of observable confounders[4]. This can make the use of standard regression analysis problematic, as the treatment and control groups may lack common support such that differences between treatment and

---

[4]Table A1 in the appendix provides a complete comparison of standardised differences between the two groups in terms of a variety of potential confounders for both raw and matched data (using model 4 below). While there is no universally agreed cut-point, standardised differences greater than 0.25 are often considered evidence of imbalance and marked with an asterisk in the table (Rubin 2001).

control would be based on extrapolation.

Matching is generally useful in situations where the observable confounders (e.g. education level, age) that (1) are related to the outcome (Brexit vote), but (2) also would have affected whether someone received the treatment (has learned to speak a foreign language), and (3) are plausibly exogenous to the 'treatment decision'. In other words, factors such as age and education affect both the Brexit vote and whether someone learned a foreign language, but neither the Brexit decision nor the reported language skills can plausibly have retroactively affected a respondent's highest level of educational qualification.

Testing the raw difference between the groups – as it is visible in Figure 3 – can be misleading, as treatment and control groups may differ in characteristics that are related to both the vote choice and foreign language skills which would confound our estimates. Absent randomisation, matching methods can still be used to estimate average treatment effects if two requirements are met: (1) conditional on observed covariates, potential outcomes are independent of treatment, i.e. the 'conditional independence assumption', and (2) there is 'common support', meaning sufficient overlap in the distributions of the treatment and control groups. This second assumption is testable – and the matched data indeed provide sufficiently similar groups, as can be seen in the balance tests in Table A3 and the overlap plot in Figure A1. Threats to conditional independence due to unobserved confounders on the other hand are invisible. That being said, the tests reported below are an attempt at a conservative estimate by matching not only on obvious, plausibly exogenous, predictors of foreign language skills such as demographics and education that are also related to the Brexit vote, but by also showing estimates for models that additionally match on strong predictors of political opinions, namely party identification and personality, that are less obviously related to language skills.

Propensity score matching first estimates a model for the propensity of being treated – in this case being conversational in a language other than English – conditional on a set of covariates. This propensity score is then used to match one or more observations from the control group to the treated observations, so that the matched control group

15

resembles the treatment group (Caliendo and Kopeinig, 2008). The difference in the referendum vote between the matched treatment and control groups is our estimate of the Average Treatment Effect on the Treated (ATT). Table 2 below provides different ATT estimates using propensity score matching with different sets of covariates.

Table 2: Matching estimates of the effect of speaking a second language on the referendum vote choice

	(1) raw t-test	(2) PSM	(3) PSM	(4) PSM
Average Treatment Effect (ATT)	13.27	5.29	3.20	3.76
Standard Error	(0.82)	(1.55)	(1.39)	(1.49)
p-value	0.001	0.001	0.021	0.011
Demographics *(age, gender, income, region, education, non-white, foreign-born parents)*		✓	✓	✓
Party identification			✓	✓
Personality factors				✓
N	20600	13167	13167	12353

The first column just shows a raw t-test between the two groups. As we saw from Figure 3 above, those speaking a language other than English are 13% more likely to have voted to remain. This is a large effect, no doubt in part due to the fact that monolinguals and multilinguals are different groups of people in a variety of respects. To address this, the estimate reported in column 2 is based on a model that uses propensity scores to match voters speaking another language to voters who are identical (or similar) in terms of education, age, gender, income, region, being non-white, and having foreign-born parents, but who crucially do not speak another language. The difference between these groups is a more credible 5% percent, and statistically significant. It means that voters who speak a foreign language are five percent more likely to have voted for Remain than a group of otherwise very similar voters with no foreign language skills.

In order to provide an additional benchmark, it is worthwhile to consider whether there are characteristics that are known to strongly determine the Brexit vote choice and that may be related to foreign language skills, and potentially indicative of other unobserved confounders. Factors such as personality factors and party identification fit this description. While this technically goes beyond what propensity score matching is

intended to be used for, it can be considered a robustness check as it should bias the treatment estimate down. The model in column 3 adds party identification to the list of covariates to match on, and column 4 also includes values of the big five personality traits[5]. The results show that even after matching respondents on personality factors and party identification, as well as demographics, there remains a three to four percent difference between those who speak another language and those who do not.

5   Limitations

In terms of limitations, the analysis presented here is constrained by the available data. Survey questions on self-reported language skills are potentially problematic in a variety of ways (Edele et al., 2015). The most obvious being misreporting, or in this case an overly generous assessment of what "conversational" means. If this is not random, but systematically related to other individual characteristics such as personality or educational background due to social desirability bias or interviewer effects, results could be biased. Furthermore, if especially Europhile, cosmopolitan or 'transnational' individuals feel a stronger need to report having foreign language skills, or over-estimate their abilities more, results would certainly be biased. It would therefore be desirable to have better questions on foreign language skills. A simple alternative would be to make the questions more concrete by giving respondents examples of situations in which they would use their languages skills, e.g. "are you able to order food/discuss politics/discuss a film in a foreign language...". A good example of a similar attempt is in Eurobarometer 65 (cited in Kuhn, 2011), which asked: "In the last 12 months, have you read a book, newspaper or magazine in a language other than your mother tongue". A more complex alternative would be to actually test language skills. Since surveys are computer-based and increasingly completed online, these could for example include a few – ideally standardised – survey items that test actual skills in commonly spoken foreign languages following the self-report question. While challenging to design and validate, this would allow for a much more thorough assessment of actual language skills. Perhaps this would

---

[5]Descriptions of all variables are in Table A1 in the appendix.

be an interesting project for linguists interested in survey research.

6 Conclusion

This chapter argued that the UK's relative lack of foreign language skills has contributed to the outcome of the Brexit referendum. Theory suggests that speaking foreign languages reduces perceptions of cultural distance and contributes to the formation of transnational identities. Research also shows a link between language skills and European identity (Kuhn, 2015; Díez Medrano, 2018). The hypothesis that voters who "speak a language other than English at conversational level" are more likely to have voted to remain in the European Union finds support. Data from the referendum wave of the British Election Study show that voters with foreign language skills overwhelmingly voted to remain. Voters with foreign languages skills are however quite different from those without in a variety of ways, most notably in terms of educational attainment, income, and family background, e.g. being foreign-born. Using matching methods, which allow for a paired comparison of otherwise similar respondents in terms of these demographic differences, it is possible to estimate the effect of foreign language skills on the referendum vote. The effect of foreign language skills – while taking into account differences in education level, age, gender, income, region, and family background in terms of being non-white or foreign-born parents – is about 5%. This means otherwise very similar voters who speak a foreign language were 5% more likely to have voted for Remain. The analysis also indicates that a significant effect of foreign language skills remains, even when taking into account additional factors such as party preference and personality differences.

The analysis presented here lends support to theories that link foreign language skills and (European) identity formation. It speaks to the literature on language and politics, showing that multilingualism can have an important influence on political behaviour, and it also corroborates general findings from the European public opinion and European identity literature in the context of the UK's highly consequential 2016 referendum on membership in the European Union.

Overall, the chapter has attempted to make three contributions to the literature on language and politics, specifically multilingualism and political identity, as well as the work on European identity, Euroscepticism and the causes of Brexit. The main contribution is the identification of the (lack of) foreign language skills as a factor that may have affected the Brexit decision. A large variety of important influences have already been discussed and analysed in the literature sparked by the Brexit referendum (cf Hobolt, 2016; Clarke et al., 2017), but the exceptionally low level of foreign language skills in Britain has so far been overlooked. Another contribution lies in the fact that the Brexit referendum provides an important test case. Much of the extant research on language, transnationalism and European identity relies on a variety of purely attitudinal measures of European identification. The Brexit vote choice was an example of highly consequential political behaviour. Finally, the question of foreign language skills goes beyond more established narratives of age, education and the economically or culturally 'left behind' as causes of Brexit. This is important because these factors such as age, education, location and personality, are difficult or impossible to influence. Language skills and foreign language learning by contrast are one of the variables that can actually be affected by government policy. In this sense, the chapter does also provide a policy implication. European Union policies in support of language learning and transnational exchange are considered important tools for creating a European identity and a cohesive union. The results presented here support this view with evidence, as the data suggest that the lack of foreign language skills played some part in the outcome of the Brexit referendum.

References

Allport, G. W. (1954), *The nature of prejudice*, Addison-Wesley.

Anderson, B. (1983), *Imagined Communities: Reflections on the Origins and Spread of Nationalism*, Verso, London.

Benet-Martínez, V., Leu, J., Lee, F. and Morris, M. (2002), 'Negotiating biculturalism: Cultural frame switching in biculturals with oppositional versus compatible cultural identities', *Journal of Cross-Cultural Psychology* **33**, 492–516.

Benet-Martínez, V. and Haritatos, J. (2005), 'Bicultural identity integration (bii): Components and psychosocial antecedents', *Journal of personality* **73**(4), 1015–1050.

Blommaert, J. and Verschueren, J. (1992), 'The role of language in european nationalist ideologies. pragmatics', *Quarterly Publication of the International Pragmatics Association (IPrA)* **2**(3), 355–375.

Breidbach, S. (2002), 'European communicative integration: the function of foreign language teaching for the development of a european public sphere', *Language Culture and Curriculum* **15**(3), 273–283.

Byram, M. (1994), *Teaching-and-learning language-and-culture (Vol. 100)*, Multilingual matters, Bristol.

Calhoun, C. (1996), Social theory and the public sphere, *in* B. Turner, ed., 'The Blackwell Companion to Social Theory', Blackwell, Oxford, pp. 429–470.

Caliendo, M. and Kopeinig, S. (2008), 'Some practical guidance for the implementation of propensity score matching', *Journal of economic surveys* **22**(1), 31–72.

Carreras, M., Irepoglu Carreras, Y. and Bowler, S. (2019), 'Long-Term Economic Distress, Cultural Backlash, and Support for Brexit.', *Comparative Political Studies* **52**(9), 1396–1424.

Clarke, H. D., Goodwin, M. and Whiteley, P. (2017), *Brexit: Why Britain Voted to Leave the European Union*, Cambridge University Press, Cambridge.

Crisp, R. J. and Turner, R. N. (2009), 'Can imagined interactions produce positive perceptions?: Reducing prejudice through simulated social contact', *American psychologist* **64**(4), 231.

Deutsch, K. W., Burrell, S. ., Kann, R. A., Lee, M., Lichterman, M., Lindgren, R., Lowenheim, F. and Van Wagenen, R. (1957), *Political Community and the North Atlantic Area*, Princeton University Press.

Díez Medrano, J. (2018), 'Multilingualism and european identification', *Sociological Inquiry* **88**(3), 410–434.

Doerr, N. (2012), 'Translating democracy: How activists in the european social forum practice multilingual deliberation', *European Political Science Review* **4**(3), 361–384.

Edele, A., Seuring, J., Kristen, C. and Stanat, P. (2015), 'Why bother with testing? the validity of immigrants' self-assessed language proficiency', *Social Science Research* **52**, 99–123.

Eder, K. and Kantner, C. (2005), 'Transnationale Resonanzstrukturen in Europa. Eine Kritik der Rede vom Öffentlichkeitsdefizit.', *Kölner Zeitschrift für Soziologie und*

Sozialpsychologie **40**, 306–331.

Eriksen, E. O. (2005), 'An emerging european public sphere', *European Journal of Social Theory* **8**(3), 341–363.

European Commission (2017), *Standard Eurobarometer 86 - Public opinion in the European Union*, European Commission, Directorate-General for Communication, Report.

Eurostat (2016), 'Foreign language skills statistics', `https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_skills_statistics`.

Fieldhouse, E., Green., J., Evans., G., Schmitt, H., van der Eijk, C., Mellon, J. and Prosser, C. (2017), 'British Election Study Internet Panel, Wave 8', `https://doi.org/10.15127/1.293723`.

Fligstein, N. (2008), *Euroclash: The EU, European identity, and the future of Europe*, Oxford University Press, Oxford.

Fraser, N. (1990), 'Rethinking the public sphere: A contribution to the critique of actually existing democracy', *Social text* (25/26), 56–80.

Gerhards, J. and Neidhardt, F. (1990), *Strukturen und Funktionen moderner Öffentlichkeit: Fragestellungen und Ansätze*, (FS III 90-101). WZB Discussion Paper.

Goodwin, M. J. and Heath, O. (2016), 'The 2016 referendum, brexit and the left behind: An aggregate-level analysis of the result', *The Political Quarterly* **87**(3), 323–332.

Habermas, J. (1974), 'The public sphere: An encyclopedia article (1964), translated by lennox, s. and lennox, f', *New German Critique, (* **3**, 49–55.

Habermas, J. (1981), *Theorie des kommunikativen Handelns*, Suhrkamp, Frankfurt.

Hobolt, S. B. (2016), 'The brexit vote: a divided nation, a divided continent', *Journal of European Public Policy* **23**(9), 1259–1277.

Hobsbawm, E. J. (1990), *Nations and nationalism since 1780: Programme, myth, reality*, Cambridge University Press.

Kappe, R. (2015), The political psychology of UKIP: Personality factors, authoritarianism and support for right-wing populist parties in Britain. *Unpublished Manuscript*, presented at the 2015 Elections, Public Opinion and Parties (EPOP) conference.

Koopmans, R. (2007), 'Who inhabits the european public sphere?', *Winners and losers, supporters and opponents in Europeanised political debates* **46**(2), 183–210.

Kuhn, T. (2011), 'Individual transnationalism, globalisation and euroscepticism: An empirical test of deutsch's transactionalist theory', *European Journal of Political Research* **50**(6), 811–837.

Kuhn, T. (2015), *Experiencing European integration: Transnational lives and European identity*, Oxford University Press, Oxford.

Mau, S., Mewes, J. and Zimmermann, A. (2008), 'Cosmopolitan attitudes through transnational social practices?', *Global Networks* **8**(1), 1–24.

Padilla, A. M. and Perez, W. (2003), 'Acculturation, social identity, and social cognition: A new perspective', *Hispanic Journal of Behavioral Sciences* **25**(1), 35–55.

Pana, A. (2015), Models of the european public sphere. Working Paper, Center for EU Communication Studies, National School of Political Studies and Public Administration (NSPSPA), Romania.

Pavlenko, A. and Blackledge, A. E. (2004), *Negotiation of identities in multilingual contexts (Vol. 45)*, Multilingual Matters, Bristol.

Pettigrew, T. F. and Tropp, L. R. (2006), 'A meta-analytic test of intergroup contact theory', *Journal of personality and social psychology* **90**(5), 751.

Risse, T. (2003), 'Toward a european public sphere?', *Theoretical Considerations, European Union Studies Association, Nashville TN: March* pp. 26–30.

Risse, T. (2015*a*), *A community of Europeans?: Transnational identities and public spheres*, Cornell University Press.

Risse, T. E. (2015*b*), *European public spheres*, Cambridge University Press.

Schecter, S. R. and Bayley, R. J. (2005), *Language as cultural practice: Mexicanos en el norte*, Routledge, Abingdon-on-Thames.

Stokes, B. (2017), 'What it takes to truly be 'one of us'', *Pew Research Center Report, February* **2017**.

Strani, K. (2010), Communicative rationality and the challenge of systems theory, *in* C. B. Grant, ed., 'Beyond universal pragmatics: studies in the philosophy of communication', Peter Lang, Oxford.

Strani, K. (2014), 'The impact of socio-political change on public sphere theory: exploring the relevance of communicative rationality', *International Journal of Cross-Cultural Studies and Environmental Communication* **3**(1), 31–40.

Trenz, H. J. and Eder, K. (2004), 'The democratizing dynamics of a european public sphere: Towards a theory of democratic functionalism', *European Journal of Social Theory* **7**(1), 5–25.

Walter, S. (2017), 'Three models of the european public sphere: An analysis of the actor structure in eu news', *Journalism studies* **18**(6), 749–770.

Wodak, R., de Cillia, R., Reisigl, M. and Liebhart, K. (2009), *Discursive construction of national identity*, Edinburgh University Press, Edinburgh.

Wright, S. (2016), *Language policy and language planning: From nationalism to globalisation*, Springer, New York.

Appendix

Table A1: Variables and operationalisation

Concept	Operationalisation	BES variable name
Multilingualism	*"Do you speak a language other than English [or Welsh] at conversational level"*.	languageSkills, languageSkillsWelsh
Referendum Vote	If you do vote in the referendum on Britain's membership of the European Union, how do you think you will vote?	euRefVote
Education	Respondent's highest education level: No Qualifications, GCSE D-G, GCSE A*-C, A-level, Undergraduate, Postgraduate	profile_education_level
Age	Respondent's age	age
Gender	Dummy variable based on question "Are you male or female?"	gender
Income	Fifteen income categories based on question "What is your gross household income?"	profile_ gross_household
Non-white background	Dummy based on question "To which of these groups do you consider you belong?"	profile_ ethnicity
Foreign born parent(s)	Dummy based on question "Were either of your parents born outside the United Kingdom?"	parentsForeign
Region	UK Government Office Region	profile_GOR_pdl
Party identification	"Generally speaking, do you think of yourself as Labour, Conservative, Liberal Democrat or what?"	partyId
Personality	Respondent's score on 0-10 scale for the "Big 5" personality factors: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism based on Ten Item Personality Inventory (TIPI)	personality_openness personality_conscientiousness personality_extraversion personality_agreeableness personality_neuroticism

Table A2: Summary statistics and correlation with multilingualism and remain vote for age and personality factors

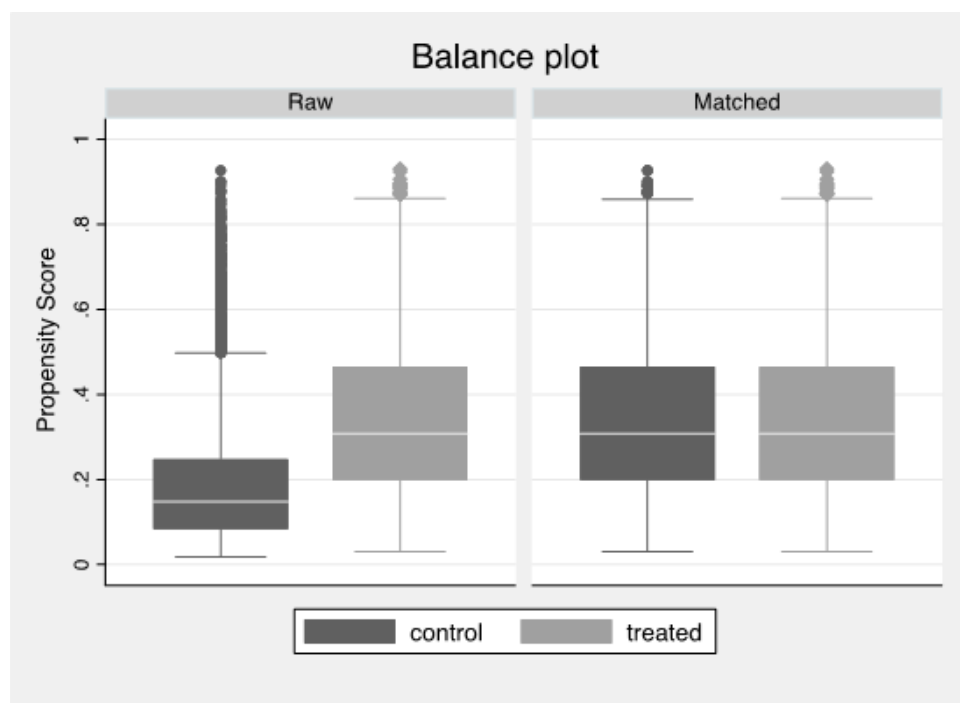| Variable | Summary Statistics | | | | Correlation with | |
	Mean	SD	Min	Max	Multilingualism	Remain Vote
Age	53.7	15.3	18	96	-0.02*	-0.21*
Agreeableness	6.09	1.77	0	10	-0.01	-0.01*
Conscientiousness	6.84	1.84	0	10	0.02*	-0.07*
Extraversion	4.09	2.18	0	10	0.08*	-0.01
Neuroticism	3.67	2.19	0	10	-0.03*	0.03*
Openness	5.52	1.71	0	10	0.14*	0.10*

Figure A1: Balance plot showing common support

Table A3: Covariate balance pre- and post-matching

	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Age	-0.008	-0.012	1.162	1.224
Gender	0.046	0.013	1.000	0.999
Income	0.239	0.001	1.141	1.016
Region				
North West	-0.058	0.060	0.855	1.202
Yorkshire and Humber	-0.048	-0.029	0.868	0.918
East Midlands	-0.065	0.028	0.803	1.112
West Midlands	-0.051	-0.044	0.853	0.871
East of England	-0.042	0.023	0.889	1.071
London	0.290*	-0.024	1.777	0.966
South East	-0.025	-0.003	0.947	0.993
South West	-0.070	-0.007	0.810	0.978
Wales	0.054	-0.011	3.003	0.858
Scotland	0.031	0.019	1.060	1.036
Education				
GCSE D-G	-0.203	-0.013	0.336	0.911
GCSE A*-C	-0.395*	0.036	0.518	1.096
A-level	-0.091	0.019	0.868	1.035
Undergraduate	0.353*	-0.053	1.206	0.994
Postgrad	0.396*	0.017	2.393	1.025
Non-white background	0.490*	0.044	3.757	1.070
Foreign born parents	0.509*	0.017	2.521	1.018
Personality				
Agreeableness	-0.016	-0.044	0.968	0.924
Conscientiousness	0.047	-0.032	1.080	1.063
Extraversion	0.175	0.035	1.107	1.027
Neuroticism	-0.063	0.033	1.018	1.059
Openness	0.331*	-0.030	1.098	1.125
Party ID				
Labour	-0.005	-0.016	0.996	0.987
Liberal Democrat	0.109	-0.009	1.411	0.976
SNP	-0.011	-0.040	0.955	0.850
UKIP	-0.142	-0.019	0.617	0.928
Green Party	0.093	-0.022	1.673	0.903
Other	0.067	0.010	1.995	1.090
No - none	-0.019	0.033	0.966	1.068
Don't know	-0.025	0.023	0.874	1.143