UCD GEARY INSTITUTE FOR PUBLIC POLICY
DISCUSSION PAPER SERIES

When one side stays home:
A joint model of turnout and vote choice

Johan A. Elkink
School of Politics and International Relations, University College Dublin

Sarah Parlane
School of Economics, University College Dublin

Thomas Sattler
Université de Genève (UniGE),
Département de Science Politique et Relations Internationales

Geary WP2020/12
November 26, 2020

When one side stays home:

A joint model of turnout and vote choice

Johan A. Elkink[a], Sarah Parlane[b], Thomas Sattler[c]

Abstract: Most existing studies of referendums examine voter turnout and their vote choice separately. Our theoretical model suggests that this is problematic. The model predicts that voters who generally prefer one of the two possible referendum outcomes, but who are relatively uncertain about the consequences of their preferred option, tend to abstain from voting. Greater uncertainty about a referendum option not only reduces its value, but also, for more "distant" voters, the value of participating. Uncertainty, thus, has a double effect: potential supporters of one referendum option are less likely to vote; and citizens who vote are less likely to support this option. We use data from the 'Brexit' vote to show how individual assessments of uncertainty about the two-options affect turnout and the vote. Our empirical analyses provide support for our theoretical model.

Keywords: referendums, turnout, uncertainty, European integration, electoral behavior and Brexit.

[a]Associate Professor, School of Politics and International Relations, University College Dublin. [*]Corresponding author, email: jos.elkink@ucd.ie.
[b]Assistant Professor, School of Economics,University College Dublin.
[c]Associate Professor, Université de Genève (UniGE), Département de Science Politique et Relations Internationales.

I. Introduction

Recent referendums in the European Union (EU) indicate that voter turnout has a substantial effect on the outcome of the vote. One example is the 'Brexit' referendum in the United Kingdom (UK) in June 2016 over the question whether Britain should leave the EU. Several analysts blamed low turnout among younger voters for lack of support for staying within the EU.[1] An even more obvious example is the repeated referendums in Ireland on the Nice Treaty in 2001/02. The Irish electorate first rejected the treaty and then approved it in a second referendum because supporters of the treaty who previously abstained decided to turn out to vote (Sinnott, 2003). We see similar patterns in other popular votes such as national elections. For example, the presidential elections of France in 2017 were likely influenced by the low turnout of voters from the political left.[2]

Although existing research has long suggested that turnout matters for referendum outcomes, the general mechanisms through which this happens remain unclear. The anecdotes above, for instance, concur with the view that a critical channel through which the outcome of a vote can be influenced is by affecting turnout (Zhang, 2018). In the context of EU referendums, the link between turnout and referendum outcomes has been highlighted repeatedly (Sinnott, 2003; Hobolt, 2009). But despite this consensus, most analyses still examine the decisions whether to vote and how to vote as separable choice problems (e.g., Hug and Schulz, 2007; Hobolt, 2009; Dür and Konstantinidis, 2013; Elkink and Sinnott, 2015). Only very few theoretical models explicitly model the connection between turnout and vote choice (e.g., Matsusaka, 1995; Feddersen and Pesendorfer, 1996).

Our theoretical model shows that an important mechanism linking turnout to vote choice is through uncertainty about the consequences of the referendum options. First, and contrary to the previous literature, greater uncertainty about the

---

1 *The Guardian*, "EU referendum: youth turnout almost twice as high as first thought", 10 July 2016, https://www.theguardian.com/politics/2016/jul/09/young-people-referendum-turnout-brexit-twice-as-high.

2 *BBC News*, "French election: Turnout sharply down in Le Pen-Macron battle", 7 May 2017, http://www.bbc.com/news/world-europe-39833831; *Independent*, "France presidential elections: As French go to polls, could abstentions prove decisive in Macron-Le Pen contest", 7 May 2017, http://www.independent.co.uk/news/world/europe/france-presidential-elections-le-penmacron-french-polls-vote-abstentions-odds-who-will-win-a7722356.html; see Galam (2017) for an analysis of the interaction between turnout and vote choice in the French elections.

.

consequences of one option biases turnout in favor of the less uncertain option. Specifically, turnout not just decreases as uncertainty increases, but it is a particular group—namely those favoring the more uncertain option—who stay home. From the point of view of citizens who are highly uncertain about their preferred choice, the distinction between the two options blurs, and voting loses its appeal. Second, and consistent with the previous literature, greater uncertainty further reduces the support for an option in the group of participating voters that is already biased against the more uncertain option. Voters are critical of options for which the consequences are uncertain and hence tend to vote for the less uncertain option. Uncertainty, thus, has a double impact: first, voters on one side stay home; and then, some of the participating voters change their vote.

We test the predictions of our model using survey data on the 'Brexit' referendum in the UK (2016). This survey is unique because it directly measures uncertainty of the two referendum options. Specifically, respondents were asked how certain they are about the consequences of remaining in the EU and of leaving the EU prior to the referendum taking place. Furthermore, questions related to the expected impact of 'Brexit' on the UK allow us to include a proxy for the difference between the status quo and the proposed position in our spatial model of voting. We are therefore able to derive an empirical model specification closely aligned with our theoretical model (cf. Signorino, 1999).

The empirical results confirm our hypotheses. They show that voters whose preferred position is closer to the more uncertain option are less likely to vote and less likely to vote for that option. As an additional implication of our theoretical model, we also find that as the two options are closer to each other, turnout declines. Not only do we provide support for our theoretical model, we also demonstrate how the overall estimated outcome of the referendum can be substantially mis-estimated when only vote choice is taken into account.

The mechanism that we describe is not unique to referendums, but can be applied to elections more generally, at least where it concerns two possible outcomes such as the second round of the French Presidential Elections. In such a context similar variation in uncertainty about the relative policy positions, or the implications of someone being elected president, can lead to similar voter dynamics and thus a similar suboptimal election outcome.

The following section provides an intuitive understanding of our theoretical argument, prior to the formal model specification in the third section. The fourth section compares our theoretical model to existing formal models that link turnout and vote choice to the aggregate referendum outcome. The fifth section tests the empirical implications of the model on the 'Brexit' referendum data. It also evaluates how much ignoring turnout leads to an underestimation of the impact of uncertainty on the vote in this referendum.

II. Uncertainty, turnout, and vote choice

The existing literature shows that uncertainty, and related personal characteristics, strongly influence both turnout and vote choice. For instance, there is a strong relationship between education (Nie et al., 1996), cognitive ability (Denny and Doyle, 2008), political awareness (Zaller, 1992), political knowledge (Elkink and Sinnott, 2015) and both participation and vote choice, and there is indeed a longstanding literature on the topic (Campbell et al., 1960; Nie et al., 1979; Zaller, 1992; Carpini and Keeter, 1996). Research also finds that more knowledgeable citizens participate in elections and EU referendums more often (Matsusaka 1995; Hobolt 2005) and that they are more likely to support European integration proposals (Hobolt 2009; Elkink and Sinnott 2015). Uncertainty about a vote option can be caused by the general complexity of the issue being voted upon, mixed messages from different sides of the debate, the intensity of the referendum campaign, lack of predictability of international reactions, or many other reasons (Farrell and Schmitt-Beck, 2003; Nicholson, 2003). Certainly, in European integration referendums, all three causes are likely to be present. In the remainder of the argument, we are agnostic about the cause of the uncertainty, and focus on its impact.

To explore the link between turnout and referendum outcomes, our formal model examines a situation in which voters are presented with two options, e.g. in a referendum or a two-candidate election. In line with the previous literature, we assume that voters favor options that are closer to their own, most preferred outcome and dislike options that have uncertain consequences. While voters can identify which option is closer to their ideal outcome, they face some uncertainty around each option's exact position. Following the announcement that a referendum will be held, a campaign starts during which relevant information is released. This information takes the form of a signal which is positively correlated with each option's exact location. With this new information, each citizen has a more precise idea about the true consequences of the two options. This additional information, however, is not always associated with a positive value-added despite the fact that it reduces the uncertainty. Indeed, some voters may realize, during the campaign, that the option closer to them is not as close as what they originally thought. From the point of view of 'extreme' voters, this implies that the distinction between the two options blurs, which dampens their willingness to vote. The greater the uncertainty surrounding their preferred option, the more likely it is that they will lose interest.

When voters decide *how to vote*, they choose the option that is closest to their most preferred outcome, provided that they have a clear idea what the consequences of this option will be. At the same time, risk-averse voters are critical of options for which the consequences are uncertain. This implies that 'centrist' voters, i.e. voters towards the middle of the policy space, are less likely to vote for the more uncertain option. Our model, thus, predicts that greater uncertainty around one option leads to both lower turnout among supporters of that option and a reduced probability of a vote in favour of that option. Unlike predictions from previous models, turnout not just decreases as uncertainty increases, but it is a particular group—namely those favoring the more uncertain option—who stay home. What matters here is not the level of uncertainty, but the relative uncertainty levels between the two referendum options. As a result, our model is applicable in a wide range of referendums, with cooperative or non-cooperative reversion points (Hobolt, 2009) and different levels and types of uncertainty (Walter et al., 2018). Uncertainty about the consequences of the two options, thus, influences the referendum outcome in two ways. First, support for the more uncertain option decreases via turnout because 'extreme' voters on one side stay home. This biases voter turnout in favor of the less uncertain option. Second, support for the more uncertain option decreases via citizens' vote choice because the 'centrist' voters change sides.

III. Theoretical model

Theoretical Background

Recent referendums on European integration have been associated with high informational demands as diverse complex issues were at stake. Information is particularly valuable in settings where ideology plays a minor role and where one option is objectively “superior", for economic or social reasons, but voters are unsure as to which is. Feddersen and Pesendorfer (1996, 1999) show that, in such a context, uninformed voters may prefer to abstain fearing that their ballot could lead to the selection of the “inferior" option. In a setting where voters have some intrinsic preferences, Larcinese (2009) shows that the fear of being pivotal leads voters to acquire costly information. Several other formal models of voting behaviour similarly built around this relation between pivotality and the level of information (Matsusaka, 1995; Ghirardato and Katz, 2002; Larcinese, 2009; Degan and Merlo, 2011; Oliveros, 2013). To address the presence of uncertainty, it is not uncommon for member states to issue pamphlets highlighting the main political, economic, and social implications associated with each outcome. This paper brings to light the impact that both, the lack and the supply of information, have on participation and voting.

We consider individuals who are guided by their intrinsic ideology and an aversion to uncertainty. The individuals’ intrinsic preferences are captured considering a spatial model of voting for popular referendums. In this model, the choice options are positioned along a one-dimensional policy space. Each individual has an ideal point which reflects the policy that this person prefers to all other policies in the policy space. The referendum gives voters a choice between two options, neither of which they can locate with certainty. Risk aversion is then captured relying on Hobolt (2009). Voters prefer the options that are closer to their ideology, but they dislike those which are poorly defined and uncertain.

In a European context, our setting captures a situation where voters are aware that one option is generally more supportive of integration than the other. However, there is uncertainty as to how much integration this option will lead to, as well as uncertainty as to how little integration the other option will lead to (Walter et al., 2018). We therefore capture the idea that there is a difference between what one person votes for and what policy is actually implemented, even when this voter’s choice coincides with the selected option. What policies will be implemented are referred to as the true possible states of the world. Clearly, while the overall policy will reflect the popular vote, its implementation (and thus its precise location) will depend on many other factors such as the government's ability to negotiate, its composition and factors affecting the economic and social context.

In relation to European integration, citizens can be uncertain for many reasons. They may be unsure how other EU members respond when the country rejects more integration or leaves the European Union. In both cases, the future relationship with the remaining EU countries needs to be renegotiated. Similarly, it is unclear how great the loss of national sovereignty is, i.e. the government's ability to freely choose domestic policies, when countries integrate more (Sattler and Urpelainen, 2012) or less. It is common that treaty provisions need to be clarified and interpreted after a new integration proposal is accepted. Even for well-informed citizens, it is difficult to anticipate how the advances or reversals of European integration will affect their country or themselves.[3]

Model setup

Consider a policy space representing the set of policy options available in a particular policy field. Without loss of generality, we assume that the policy space varies between 0 and 1. Applied to EU referendums, the policy space reflects the depth of European integration. A value of 0 represents a situation, in which European integration ends, and all member states continue as individual states. A value of 1 represents a situation where Europe fully integrates into a federal state. Voters are distributed across the policy space according to a probability distribution $F(.)$. Each voter has an ideal point $i \in [0,1]$, which, in European politics, reflects the depth of European integration that the voter favors most. A greater i represent more pro-integrationist EU attitudes.

In the referendum, citizens have the choice between two options with different (expected) positions in the policy space: t_0 and t_1, $0 < t_0 \leq t_1 < 1$. In the 'Brexit' referendum, for instance, option t_1 represents the status quo, which is the current level of European integration; and option t_0 represents the proposal to leave the European Union, which reduces the current integration level. In other, past integration referendums, such as those on the Maastricht, Nice and Lisbon Treaties, the situation is

[3] We depart from the assumption in the existing literature that only the new treaty is associated with uncertainty. By assuming that both options can be uncertain, our model is more general and thus subsumes the existing models.

the reverse: option t_0 represents the status quo, and option t_1 represents the proposed treaty.

The citizen's decision problem is complicated by uncertainty about the consequences of the two options. In our model, uncertainty about an option means that the outcomes t_0 and t_1 are random variables. Their realization captures the level of integration that will be implemented which is not perfectly predictable. We assume that there are two possible true positions for each option. Specifically, t_1 is either at a low (L) position, $t - \Delta_1$, in which case the proposed option really leads to a moderate deepening of EU integration, or at a high (H) position, $t + \Delta_1$, in which case the proposed option really leads to significantly more EU integration. Similarly, t_0 is either at , $\alpha t - \Delta_0$ or $\alpha t + \Delta_0$. Furthermore we assume that the joint distribution is such that there are only two possible states of the world: *LL* = ($\alpha t - \Delta_0$, $t - \Delta_1$,) and *HH*=($\alpha t + \Delta_0$, $t + \Delta_1$,).[4] One can motivate the fact that we consider only two states of the world considering that some of the factors that affect implementation are not dependent on which option ends up being selected. The variables $t, \alpha t, \Delta_0$ and Δ_1 are common knowledge and such that

$$\alpha t + \Delta_0 \leq t - \Delta_1 \text{ and } 0 < \alpha t - \Delta_0 \leq t + \Delta_1 < 1.$$

In words, this means that voters know that t_1 leads to greater integration than t_0 and that the policy space captures all eventualities.

To keep matters simple, we consider that, *a priori*, each state of the world is equally likely: $\Pr(LL) = \Pr(HH) = \frac{1}{2}$. It follows that, *a priori*, the expectations of t_0 and t_1 are $E(t_0) = \alpha t$ and $E(t_1) = t$, while their variances are $V(t_w) = \Delta_w^2, w = 0,1$. Hence the uncertainty about an option is reflected by the size of Δ_w. When Δ_w increases, citizens are more uncertain about the true consequences of option $t_w, w = 0,1$. This allows us to explore what happens when one of these options is associated with more, or less uncertainty by increasing or decreasing Δ_0 or Δ_1 in the analysis below.

Citizens are risk averse and dislike uncertainty. To capture this feature, we follow Hobolt (2009) whereby individuals have a concave, quadratic utility function. This is the

---

[4] We have analysed the results when the states ($\alpha t - \Delta_0$, $t + \Delta_1$) and ($\alpha t - \Delta_0$, $t + \Delta_1$) are added to the preceding two. This extension shows the merits of keeping the model simple.

conventional approach to modelling risk aversion in utility functions.[5] Specifically, the utility function of a voter located at i is based on voters minimizing the squared distance from their ideal point:

$$U_i(t_w) = -(t_w - i)^2, \quad (1)$$

where $w \in \{0,1\}$ for the two referendum options. Given this specification, the expected utility about a particular outcome is given by the following expression:

$$E[U_i(t_w)] = -(E(t_w) - i)^2 - V(t_w). \quad (2)$$

Because we use a squared distance in the utility function (1), the variance has a negative impact on the utility—the more uncertain a voter is about an option, the less likely she is to vote for that option.

Figure 1, below, illustrates this setup graphically. Euroskeptic citizens i_1 and i_2 are closer to option t_0 than to option t_1 and, hence, should reject option t_1. Eurofriendly citizens i_4 and i_5 are closer to option t_1 and, hence, should support this option.

[5] See Hobolt (2009), ch. 2, esp. p. 49 and fn. 12.

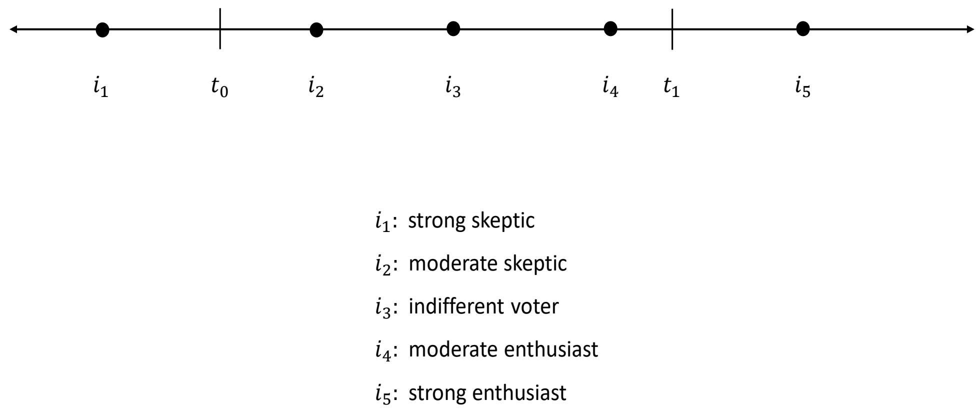

Figure 1: Possible attitudes towards European integration

Following the announcement that a referendum will be held, a campaign is held and all citizens receive a signal $s = L$ or $s = H$, which informs them about the possible consequences of a positive or negative outcome in the referendum. The signal is non-negatively correlated with the true position of the options, in line with existing strategic literature on information transmission and learning (Schneider and Cederman 1994; Schneider and Weitsman 1996; Hug 2002, ch.4; Dür and Konstantinidis 2013).[6] We consider that, *a priori*, each signal is equally likely, so that $\Pr(s = L) = \Pr(s = H) = \frac{1}{2}$.

6 In our setup, all citizens receive the same signal. It could be argued, based on "echo chamber effects" of communication, that citizens are more likely to receive signals that are more comforting given their own position (see, for instance Bessi et al., 2015; Del Vicario et al., 2016). Appendix **??** outlines how our theoretical model would change.

The joint probabilities are defined as follows:

$$\Pr(HH \cap s = H) = \Pr(LL \cap s = L) = p,$$
$$\Pr(HH \cap s = L) = \Pr(LL \cap s = H) = \frac{1}{2} - p.$$

The variable p varies between ¼ and ½ . This joint probability is a measure of accuracy or level of informativeness of the received signal. When $p = \frac{1}{2}$ the signal is fully informative about the true position of an option w. When $p = \frac{1}{4}$ the signal is not informative at all, citizens learn nothing about the true position of an option w.

We can compute the conditional probabilities associated with each possible state of the world using the unconditional and joint probabilities in the previous paragraph. This conditional probability represents the probability that the true position of a referendum option is either high (HH) or low (LL), given that the citizen receives a high or low signal. Following Bayes' Theorem and given that either signal is equally likely, we have

$$\Pr(HH|s = H) = \Pr(LL|s = L) = 2p,$$

while

$$\Pr(HH|s = L) = \Pr(LL|s = H) = 1 - 2p.$$

Notice that, within this setting, we also have

$$\Pr(t_w = H|s = H) = \Pr(t_w = L|s = L) = 2p.$$

Therefore, if $p = \frac{1}{2}$, the signal perfectly correlates with the actual position, and, once the signal is received, the voter has perfect knowledge of the position of both referendum options. If $p = \frac{1}{4}$ the opposite occurs, and the voter has as little information after receiving the signal as she had beforehand. Put differently, the conditional probability of an outcome after receiving the signal is identical to the unconditional probability of the outcome before receiving the signal.[7] This means that during the campaign, the variance can decrease, which reduces uncertainty about the absolute and relative positions of the status quo and the treaty proposal.

The timing that all citizens face is as follows. First, at time T = 0, the government announces a referendum. Citizens can vote for either option. Citizens are free to vote. At time T = 1, a campaign starts, and all citizens receive some information correlated with the true outcomes (the signal). Finally, at time T = 2, citizens who vote decide between t_0 and t_1 and cast their vote. As is customary in such sequential settings, we assume that citizens can perfectly anticipate the outcome of the subsequent period.

Voting in favor or against

The voter would vote in favor of option t_1 if and only if

$$E(U_i(t_1)|s) \geq E(U_i(t_0)|s), \tag{3}$$

where s is the signal the voter received during the campaign. By plugging (2) into (3) and solving for i, we can derive a threshold value, i_s, that separates voters into supporters and opponents of the two options after receiving signal s:

$$i_s = \frac{V(t_1|s) - V(t_0|s)}{2[E(t_1|s) - E(t_0|s)]} + \frac{1}{2}[E(t_1|s) + E(t_0|s)]. \tag{4}$$

Thus, if a voter is to the right of this position $(i > i_s)$ she votes in favor of the option proposing more integration, while otherwise she votes in favor of the option proposing less integration.[8] The expression for i_s implies that if there is greater uncertainty around the implications of one option than there is around the other, voters are less likely to vote in favor of the former. The literature typically assumes that the consequences of new integration proposals are more uncertain, given the highly complex nature of most European integration treaties. Similarly in the case of 'Brexit', the greater uncertainty is likely to be around the UK leaving the EU, not the status quo. However, the

---

[7] To illustrate this differently, obtaining the information reduces the variance to $V(t_\omega) = 8p\Delta^2_\omega(1 - 2p)$, such that if p = $\frac{1}{4}$, the variance remains the same and if p = $\frac{1}{2}$, the variance is zero.

[8] Strictly speaking, we might consider a voter deciding not to vote, after investing in information acquisition, if $i = i_s$. This is, however, an infinitely small probability.

reverse can also be imagined. For example, in Ireland in the Lisbon II referendum, a significant amount of rhetoric revolved around the uncertain outcome if Ireland would not support the treaty, given the dependency of the country on financial aid from European partners. In such a scenario, the uncertainty around the status quo might be greater than the uncertainty around the treaty, leading to increased support for the treaty. Using our setup, the indifferent types for *s* = *L* and *s* = *H* are given by

$$i_s = \frac{t^2(1-\alpha^2) + (\Delta_1^2 - \Delta_0^2) + 2\phi t(4p-1)(\Delta_1 - \alpha\Delta_0)}{2t(1-\alpha) + 2\phi(4p-1)(\Delta_1 - \Delta_0)}, \quad (5)$$

where $\phi = -1$ when $s = L$ and $\phi = 1$ when $s = H$. Clearly, the division of voters is impacted by the signal received (via ϕ), the precision of the signal (via p) and the variance associated with each option. To test our hypothesis, we can simplify matters and eliminate the noise added by the signal and its precision.[9] Considering that signals are uninformative allows us to focus on the role played by the variances. When $p = \frac{1}{4}$, a voter supports the treaty (conditional on voting) provided her position is such that

$$2ti(1-\alpha) \geq t^2(1-\alpha^2) + (\Delta_1^2 - \Delta_0^2). \quad (6)$$

The following hypothesis captures the main observable implication from this discussion:

H_1: Voters who participate in the referendum are more likely to vote against the more uncertain option.

Participation

It is undeniable that, in real life, motives for participation diverge across citizens. It is also demonstrated that political campaigns can have an impact on participation (Farrell and SchmittBeck, 2003). Indeed, one objective of the political campaigns consists in informing citizens about the issues at stake to incentivise them to participate. Kenski and Stroud (2006) and, more recently, Dimitrova et al. (2014)

[9] To capture the impact of the signal's precision, note that $i_s = \frac{1}{2}t(1+\alpha) + \phi\Delta(4p-1)$ when $\Delta_0 = \Delta_1 = \Delta$. Thus, an increase in the precision of the signal widens the gap between i_L and i_H.

focus on digital information and provide evidence of the positive impact that such information can have on participation.

While the vote itself is an expression of an individual's preference, we capture the decision to participate and engage in the referendum as a reflection of the value-added by the information available during the campaign. To do so we contrast the value given to the potential outcomes prior to acquiring any information with the value given to the same outcomes once the campaign is over. We then postulate that a citizen decides to vote if the value she attributes to the referendum's outcomes has increased during the campaign.

The signal received during the campaign reduces the variance of options t_0 and t_1. This has an uncontroversial positive impact on the expected utility. Each signal, however, has another impact on the expected utility which may be positive or negative. For voters located on the extreme left, $s = H$ informs them that both options are further away from their ideal option than what they originally thought. The opposite holds for voters located at the extreme right when they receive $s = L$. This reduces their incentive to participate as both options become closer substitutes. For voters in the middle, the campaign has a more balanced impact. They learn that one of the two options is closer to their ideal outcome while the other is further away. They are less likely to lose interest.

To be consistent with the conventional assumptions made in game theory we must consider that voters are sophisticated. In particular, voters can perfectly anticipate the signals that they can receive, and the information associated with each signal. However, to emphasize the role played by the campaign, which is otherwise limited to drawing a high or low signal, we can consider the situation prior to obtaining the signal as equivalent to a situation with a non-informative signal. More specifically, we consider that initially (prior to the campaign) voters receive a signal that is not informative (so that $p = \frac{1}{4}$ in our specification), while $p \geq \frac{1}{4}$ after the campaign. Using this information, we now calculate the expected utility of voters before and after the campaign and identify the set of voters who participate (on expectation).

Let ρ denote the probability that option t_1 is adopted following the referendum and let ρ_s denote the same probability when voters received signal $s \in \{H, L\}$. Let $U_{i,R}$ refer to a citizen's reservation utility when located at $i \in [0,1]$. This reflects the value given to the referendum's outcomes prior to the campaign when the signal is not informative:

$$U_{i,R} = \rho E[U_i(t_1)] + (1-\rho)E[U_i(t_0)], \quad (7)$$

where ρ is consistent with the fact that, *a priori*, the signal is equally likely to be high or low so that:

$$\rho = \frac{1}{2}(\rho_H + \rho_L). \quad (8)$$

Let $U_{i,P}$ be the value given to the referendum's outcomes by an informed citizen who received signal $s \in \{H, L\}$:

$$U_{i,P} = \rho_s E[U_i(t_1)|s] + (1-\rho_s)E[U_i(t_0)|s]. \quad (9)$$

Participation is measured, on expectation, by the set of voters V defined as

$$V = \{i: U_{i,R} \leq E_s(U_{i,P})\}. \quad (10)$$

Given our setting, the set of voters is defined (*ex-ante*) by all i such that

$$\sum_{s\in\{H,L\}} \rho_s\left[E(U_i(t_1)|s) - E\big(U_i(t_1)\big)\right] + (1-\rho_s)\left[E(U_i(t_0)|s) - E\big(U_i(t_0)\big)\right] \geq 0.$$

Using the expression for the utility, we calculate the above and find that a citizen located at i participates provided

$$2iA - B \geq 0, \quad (11)$$

where

$$A = (4p-1)(\rho_H - \rho_L)(\Delta_1 - \Delta_0)$$

and

$$B = 2t(4p-1)(\rho_H - \rho_L)(\Delta_1 - \alpha\Delta_0).$$

The fact that expression on the left-hand side of (11) is linear in i means that the ones who choose not to participate in the vote, if any, will be at one extremity.

Proposition: *There are three possible equilibrium outcomes. In all of these, and under consistent beliefs, option t_1 is more likely to be adopted when $s = L$ so that $\rho_H \leq \rho_L$. Participation is such that either all voters participate, or else such that some extreme voters abstain as depicted in Figure 2.*

Proof: See Appendix.

The set *V* contains all i such that

$$(\rho_H - \rho_L)[i(\Delta_1 - \Delta_0) - t(\Delta_1 - \alpha\Delta_0)] \geq 0. \quad (12)$$

As shown in the proof, whether Δ_0 or Δ_1 is higher, we always have $(\rho_H - \rho_L) < 0$. Using this in (12), we can rewrite the set of voters V as containing all i such that:

$$i(\Delta_0 - \Delta_1) - t(\alpha\Delta_0 - \Delta_1) \geq 0. \quad (13)$$

Figure 2, below, captures our findings in terms of participation.

Let us consider more carefully the circumstances under which some citizens stay at home. These are the citizens located on the extreme left when $\Delta_1 < \alpha\Delta_0$ and they are those located on the extreme right when $\Delta_1 > \frac{1-\alpha t}{1-t}\Delta_0$ (see Footnote 11). This result shows that freely available information that reduces the uncertainty is not valued equally by all citizens, despite their aversion to risk. The value added depends on the *relative* variances associated with each outcome. When the option that is closest to one voter is associated with high uncertainty relative to her less favored option, this voter can lose interest in the referendum during the campaign if she realizes that both options are further away from her. While one option offers an outcome close to her political ideology, its associated uncertainty reduces its appeal.

The following hypothesis captures the main observable implication from this discussion:

H_2: Voters at the extreme end of the policy space on the side of the more uncertain option are more likely to abstain.

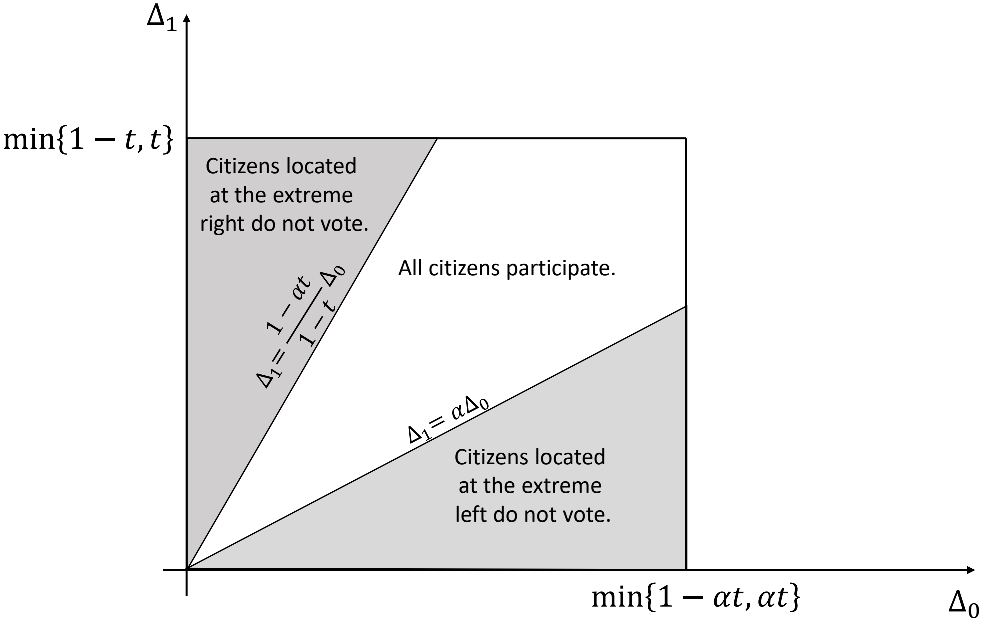

Figure 2: Graphical depiction of the theoretical model of voter turnout. The boundaries for Δ_0 and Δ_1 are such that $\alpha t - \Delta_0 > 0$, $\alpha t + \Delta_0 < 1$, $t - \Delta_1 > 0$, and $t + \Delta_1 < 1$.

IV. Discussion

A main advantage of our modelling approach is that it does not rely on the assumption that a voter fears the likelihood of being pivotal. Pivotality mostly matters in close elections (Bursztyn et al., 2018), especially when uncertainty about the election outcome (rather than the consequences of the election options) is large (Myatt, 2015). But pivotality is clearly not the main driver of turnout, and there is only weak evidence that the voter is really assessing the probability of being pivotal (Grant and Toma, 2008). In our model, pivotality is not important for turnout.

Instead, turnout varies because of the aversion towards uncertainty about the consequences of an option, which translates into diverging utilities of participation and vote choice.

In addition, our model does not require a countervailing cost component that motivates voters to turn out, as is the case in the main alternatives to the pivotal voter model. In ethical turnout models, for instance, voters face a cost of abstaining, which arises from social pressure to vote and the fear that like-minded voters learn about the voter's failure to turn out to vote for their preferred party (DellaVigna et al., 2017). In this class of models, greater turnout can affect the electoral advantage of large or small parties and in this way produce a bias. But, unlike our model, this advantage is unrelated to the ideological position of the party or referendum option (Levine and Mattozzi, 2018).

We also deviate from the assumption that there is a right and a wrong answer implying that there is an optimal decision for all. In previous models, the link between turnout and choice arises when uninformed voters decide, out of strategic considerations, not to vote. As a result, informed voters will make the decisions, and the outcome will be more optimal from the perspective of the uninformed voter (Feddersen and Pesendorfer, 1996, 1999). But in many instances, there is no objectively right or wrong answer to a referendum question. We, therefore, assume that the utilities differ between voters—either based on political ideology or personal circumstances—and that the vote decision and the importance of additional information varies between voters.

Relatedly, empirical research evaluates the conditions under which elections deliver on information aggregation—on the extent to which overall aggregate results match the underlying distribution of opinion in the public in a context of uninformed voters (Battaglini et al., 2008; Kim and Fey, 2007; Lau and Redlawsk, 1997, 2001). As far as information and uncertainty about the consequences of referendum options are related, our model suggests that the outcome of an election is not the same when voters are informed or uninformed. Especially when there is less information (and hence greater uncertainty) about one option, the referendum outcome should be biased and not reflect the underlying distribution of opinions.

V. Empirical analysis

The 'Brexit' referendum

We test our model using survey data that was collected before and after the 'Brexit' referendum in the United Kingdom in June 2016. This referendum is not only interesting because of its great political importance, but it also provides a hard test for our theoretical model. Due to the high political salience, turnout is this referendum was relatively high for reasons that are complementary to our analysis. Under these circumstances, the importance of uncertainty for turnout moves into the background because other factors, like political duty and interests, played a decisive role for the decision to vote.

The 'Brexit' referendum survey is also useful for our analysis because it differs in one important respect from most surveys after European integration referendums: the panel of referendum surveys before and after the 'Brexit' referendum asked explicitly about the uncertainty concerning the referendum proposal (leave) and the status quo (remain).[10] The question asked "How sure are you about what would happen to the UK if it left the EU or if it remained in the EU?", scoring each of the two questions on a four point scale from "very unsure" to "very sure".

Aside from the uncertainty variable, the key independent variable is the level of Euroscepticism or pro-European attitude i.[11] In the same pre-referendum survey, the question "Some say European unification should be pushed further. Others say it has already gone too far. What is your opinion?" was posed, with the answer captured on an eleven-point scale from "unification has already gone too far" to "unification should be pushed further". We rescale this to 0-1 to align with our theoretical model specification of the continuum of European integration.

---

[10] We make use of the 2014–2018 British Election Study Internet Panel (Fieldhouse et al., 2015), using variables from waves 8 and 9. In total 33,502 respondents took Wave 8, which was conducted by YouGov between 6th May 2016 and 22nd June 2016. In total 30,036 respondents took Wave 9—27,555 of these also took Wave 8, an overall wave on wave retention rate of 82.2%. Wave 9 was conducted by YouGov between 24th June 2016 and 4th July 2016, right after the referendum on 23rd June 2016..

[11] For brevity sake, we take this as a unidimensional variable, in line with the theoretical model. See Boomgaardenet al. (2011) and de Vreese et al. (2019) for a discussion on the multidimensional nature of EU support.

We proxy $\left(E(t_1) - E(t_0)\right)$ using the question in the same wave: "How much impact do you think that Britain leaving the EU would have on the country as a whole?" This variable is measured on a five-point scale from "no impact at all" to "a very large impact", which we rescale the 0-1, the theoretically maximum distance between t_0 and t_1. The combination of these variables allows us to directly translate the theoretical model into a set of regression equations that can be statistically tested.

Our analysis complements the studies that examine the main drivers of the 'Brexit' referendum. Consistent findings in this literature align with what we broadly capture in our model with pro-integration attitude (i). Whether the analysis is performed at aggregate, regional or individual level, the 'leave' vote is primarily driven by demographic factors such as age, education, and income (Hobolt, 2016; Vasilopoulou, 2016; Clarke et al., 2017; Goodwin and Heath, 2016; Goodwin and Milazzo, 2017; Matti and Zhou, 2017; Arnorsson and Zoega, 2018; Zhang, 2018). These factors largely correspond to theories of European integration related to the idea of "winners" and "losers" of European integration (Gabel, 1998; Tucker et al., 2002). Indeed, regions that are most strongly affected economically by globalisation in recent times were most likely to vote for 'Brexit' (Becker et al., 2017; Colantone and Stanig, 2018). Attitudes towards immigration and European integration also consistently pop up as important factors (Hobolt, 2016; Vasilopoulou, 2016; Clarke et al., 2017; Goodwin and Milazzo, 2017). Our analysis takes these studies as a starting point and examines how uncertainty about the referendum options factors into the political considerations of voters.

Based on this literature, as well as earlier literature on referendum voting behavior, we control for education, income, age, and gender. Table 1, below, provides summary statistics for all relevant variables.[12]

[12] *Pro integration* is a 0-1 scale of pro integration attitudes; the *Uncertainty* variables 0-1 scales of subjective assessments of uncertainty regarding the two possible outcomes, remain and leave, respectively; *Education* is a dummy variable whether the respondent has a university degree; the *Female* variable is a dummy variable; *Age* represents age in years; *Income* is an ordinal scale of income at fifteen levels; *Impact 'Brexit'* is a subjective assessment of the impact of the 'Brexit' referendum on the country as a whole.

	Minimum	Maximum	Mean	Standard deviation	N obs.
Turnout	0	1	94%		29,936
Remain vote	0	1	50%		28,069
Pro-Integration	0	1	0.32	0.31	25,847
Uncertainty (Δ_1)	0	1	0.37	0.24	23,100
Uncertainty (Δ_0)	0	1	0.44	0.25	22,960
Impact "Brexit" ($t_1 - t_0$)	0	1	0.66	0.26	25,544
Political knowledge	0	6	4.25	1.83	25,749
Efficacy	0	5	2.19	1.44	27,960
Education	0	1	48%		29,476
Income	1	15	6.70	3.52	18,615
Female	0	1	51%		30,036
Age	15	94	50.8	15.8	30,036

Table 1: Summary statistics for all variables. Mean with standard deviation for continuous variables and percentage of ones for dummy variables. Note that Education, in addition to the 48% ones, 12% were coded as ½ . For "Brexit" referendum, t =0 is the position of leaving the EU, t = 1 the status quo.

While there are many variables included in studies on voter turnout (Smets and van Ham, 2015), these are the main variables that can be expected to affect both levels of uncertainty and turnout and vote choice. While controlling for second-order effects (Reif and Schmitt, 1980) is common in the referendum literature, none of the main parties was clearly supportive of Brexit and all parties were internally split—we therefore have no reason to expect significant differences when controlling for party support.[13] In her more general explanation of vote choice in the 'Brexit' referendum, Hobolt (2016) includes a range of other variables in the empirical analysis, including identity, trust, affiliation, and attitude variables, but these are unlikely to be confounders for the relationship between uncertainty and voting behavior.

[13] Including a dummy variable for support for an opposition party does not affect the estimates significantly. Estimates available upon request.

Analysis and results

A full test of our model requires multiple regression analysis to include appropriate controls and to incorporate the interactions between the components of our model as implied by the theoretical model. Because of the two-step nature of the decision to vote—first whether or not to participate and then how to vote—we use a Heckman selection model. For both choice problems, we use a probit Heckman model to estimate

$$\Pr(y_k = 1) = \Pr[y_k^* \geq 0]$$

with $k = 1$ for the turnout model and $k = 2$ for the vote choice model, and where y_k^* is a latent utility that follows from the theoretical model above. In the selection component of the Heckman model, the dependent variable y_1 takes the value 1 if a voter turns out to vote and 0 if not. In the outcome component, the dependent variable y_2 takes the value 1 if a voter supports the integrationist position, i.e. she votes 'remain', and 0 if she rejects integration, i.e. she votes 'leave'.[14]

Our model specification is directly derived from the theoretical model (see Appendix for details). The latent regression for participation is[15]

$$y_1^* = \beta_0 + \beta_1 i(\Delta_0 - \Delta_1) + \beta_2 \Delta_0 + \beta_3 \Delta_1 + \beta_4 i + x_1' \beta_5 + \epsilon_1, \qquad (14)$$

where x_1 is a vector of control variables. This is equivalent to equation (13). In our empirical estimation, we leverage survey questions assessing i, Δ_0 and Δ_1 directly, while we take $E[t_0]$ and $E[t_1]$ to be incorporated in the regression coefficients for the Δ_0 and Δ_1 dependent variables, respectively. It should be noted that in our theoretical model, Δ_0 and Δ_1 do not vary by voter. We take the individual answers to the survey

[14] No survey weighting was applied in the analysis. Survey weighting is primarily important when focusing on descriptive statistics or when correcting for explicit sampling schemes. Here we use multiple regression models, while the respondents are not purposely sampled based on the relevant outcome variables in the model. Nevertheless, there is potentially a correlation between participating in the referendum and participating in the survey, and we do indeed oversample voters. Endogenous sampling is a motivation to apply survey weights (Solon et al., 2015). Upon request we can provide results when applying weights based on turnout, using a linear probability model for the vote choice. See table A3 in Appendix for results when applying weights based on turnout, using a linear probability model for the vote choice.

[15] For the sake of readability, subscripts indicating the respondent have been omitted.

questions as individual assessments of this aggregate level of uncertainty. Note that since $0 < t_0 \leq t_1 < 1$we can assume $E[t_0]$ and $E[t_1]$ to be positive and therefore provide expectations with regards to the signs of the coefficient estimates. Indeed, our theoretical model suggests that $\beta_1 > 0, \beta_2 < 0$ and $\beta_3 > 0$.

The latent regression for vote choice is

$$y_2^* = \gamma_0 + \gamma_1(\Delta_0^2 - \Delta_1^2) + \gamma_2(E[t_1] - E[t_0])i + \gamma_3 i \\ + \gamma_4(E[t_1] - E[t_0]) + x_2'\gamma_5 + \epsilon_2, \tag{15}$$

where x_2 is a vector of control variables.

This empirical specification follows directly from equation (6). Since we assume that voters are aware that $t_1 > t_0$, we can take $(E[t_1] - E[t_0])$ to be the equivalent of $|E[t_1] - E[t_0]|$ and therefore use the question on the impact of the referendum on the UK as a measure of this difference. Table 2 presents the main result from our model.[16]

We first turn to the selection mechanism of the model—the turnout equation. The derived model specification suggests that there should be a positive impact of the interaction between ideal point i and the difference in uncertainty between the 'Brexit' option and remaining in the EU, respectively, $(\Delta_0 - \Delta_1)$. With or without control variables, we find indeed a strong, positive effect. This provides direct support for H_2, which summarises this interaction, stating that for pro-integration voters (high i), greater uncertainty around remaining in the EU (Δ_1) should lead to lower turnout, and that for anti-integration voters (low i), greater uncertainty around leaving the EU (Δ_0) should lead to lower turnout. Looking separately at the impact of uncertainty, we find the expected negative effect for β_2, but not the positive sign on β_3.

[16] We make use of the Toomet and Henningsen (2008) package in R for our estimations.

	(1)		(2)	
	turnout	remain	turnout	remain
Pro integration (I)	−0.459***	3.411***	−0.339***	3.445***
	(0.053)	(0.133)	(0.075)	(0.167)
Uncertainty Remain (Δ_1)	−0.193*		−0.204	
	(0.106)		(0.158)	
Uncertainty Leave (Δ_0)	−0.379***		−0.244	
	(0.108)		(0.162)	
$I \times (\Delta_0 - \Delta_1)$	0.542***		0.703**	
	(0.179)		(0.270)	
Impact Brexit ($t_1 - t0$)		1.338***		1.397***
		(0.076)		(0.110)
$(t_1 - t_0) \times I$		1.597***		1.415***
		(0.169)		(0.220)
$(\Delta_0^2 - \Delta_1^2)$		1.287***		1.216***
		(0.058)		(0.081)
N	21,607		13,388	

*p < .1; **p < .05; ***p < .01

Table 2: Heckman selection models for voting behavior in the ‘Brexit’ referendum. Both the selection mechanism, turnout, and the outcome, voting for remain, are modelled using probit regression specifications. The turnout and vote choice equations are estimated jointly using Maximum Likelihood estimation. Model 2 includes controls for education, gender, age, and income, which are omitted from the table.

Secondly, we investigate the outcome component of the model—the vote choice equation. Hypothesis H_1 states that the greater the uncertainty around ‘Brexit’, relative to the uncertainty around the status quo, the greater the probability of a vote in support of remaining in the EU, and vice versa. Here we look at the model explaining vote choice, in this case a probit regression explaining a vote in support of remaining. Following our derivation of the theoretical model, we should find that there is a positive interaction between ideal point i and the difference betweenthe status quo t_1 and leaving the EU t_0 and we indeed find a strong, positive coefficient, with or without control variables. Secondly, we expect that greater uncertainty around the 'Brexit' option should lead to more support for the option to remain in the EU, i.e. we should obtain a positive coefficient on $(\Delta_0^2 - \Delta_1^2)$, which we do. We therefore find strong support for our theoretical model, and H_1, based on the regression analysis.

Given the outcome of the referendum, one expects a model that shows how turnout and uncertainty generated the 'Brexit' outcome. Instead, our empirical analysis shows how uncertainty and turnout reduced support for 'Brexit'. The key insight from this analysis is that British voters are sufficiently skeptic of European integration that had they been less uncertain about the consequences of the 'Brexit' option, the result would have been even more strongly in the 'Brexit' direction. High levels of uncertainty around the 'Brexit' option reduced turnout and made some moderate supporters reluctant to vote leave.

It should be pointed out that the interpretation of an interaction effect in a binary choice model is notoriously problematic (Ai and Norton, 2003; Berry et al., 2010), due to the inherent interactions between all exogenous variables in a non-linear model. Ai and Norton (2003) demonstrate how a positive and significant coefficient on an interaction term is neither a sufficient nor a necessary requirement for a positive interaction effect. Indeed, in our turnout model specification we find a positive and significant interaction effect when we use a non-linear logistic or probit regression, but we find a negative and significant interaction effect when we use a linear probability model.[17] It is important to point out, however, that Ai and Norton (2003) focus on the expected value of the outcome variable, thus the impact of the exogenous variables on the probability of observing a positive outcome. Our theoretical model, however, does not stipulate that the probability (y), but that the underlying utility function (y^*) is affected by the interaction term. In this case we can directly interpret the coefficient in the probit model (Berry et al., 2010, 261).

[17] These empirical results are available upon request.

When one side stays home

We have demonstrated how uncertainty, partly through turnout, resulted in weaker support for 'Brexit' than might have been, given British attitudes towards integration. This also shows the importance of incorporating turnout in empirical models explaining referendum voting behavior. Here we will evaluate the extent to which ignoring turnout biases aggregate empirical predictions of the referendum vote.

To do so we estimate predicted vote probabilities for a hypothetical set of voters, who are at the median of all variables, but vary on pro-integration attitudes and (relative) levels of uncertainty. For these we calculate predicted probabilities to vote in favor of remaining within the EU based on only the outcome component of the Heckman model. We then calculate predicted participation probabilities based on the selection component multiplied by the probability to vote 'remain'. Figure 3, on the left, shows the difference in predicted probabilities when only looking at the vote choice model, or when including turnout probabilities—i.e. by calculating $P(\textit{remain vote}) = P(\textit{remain vote} \mid \textit{participation}) \cdot P(\textit{participation})$.[18] This can lead to an overestimate of the support for the 'remain' option by up to approximately 20%. When looking at the more common situation of slightly greater uncertainty around 'Brexit' than around the status quo, we still overestimate the 'remain' vote by 8%. Given that the 'Brexit' option won by 3.8 percentage points, a prediction error of 8 percentage points is substantial.

An alternative approach is not to take a hypothetical population keeping variables at their median, but to use the sample distribution of the survey to calculate predicted probabilities. A similar plot can then be produced, which is available on the right in Figure 3 below. Here the misestimation is even larger. For those voters who are very uncertain about remaining within the EU, but quite confident about the situation in the UK after 'Brexit', we dramatically overestimate support for 'remain'. This is of course a rare combination—26% of the respondents are more uncertain about 'Brexit' than about the status quo, while the reverse is true for only 9%. Based on the empirical distribution of covariates, we therefore expect that the overall support for the 'remain' option in the referendum is overestimated by up to 7 percentages points when only vote

[18] Curves are smoothened curves using the generalized additive model implementation of the geom_smooth() function in Version 2.2.1 of the ggplot2 library in R (Wickham, 2015), using second order polynomials.

choice and not turnout is taken into account. On average, across all voters in the sample, the overestimate is 2.7%.

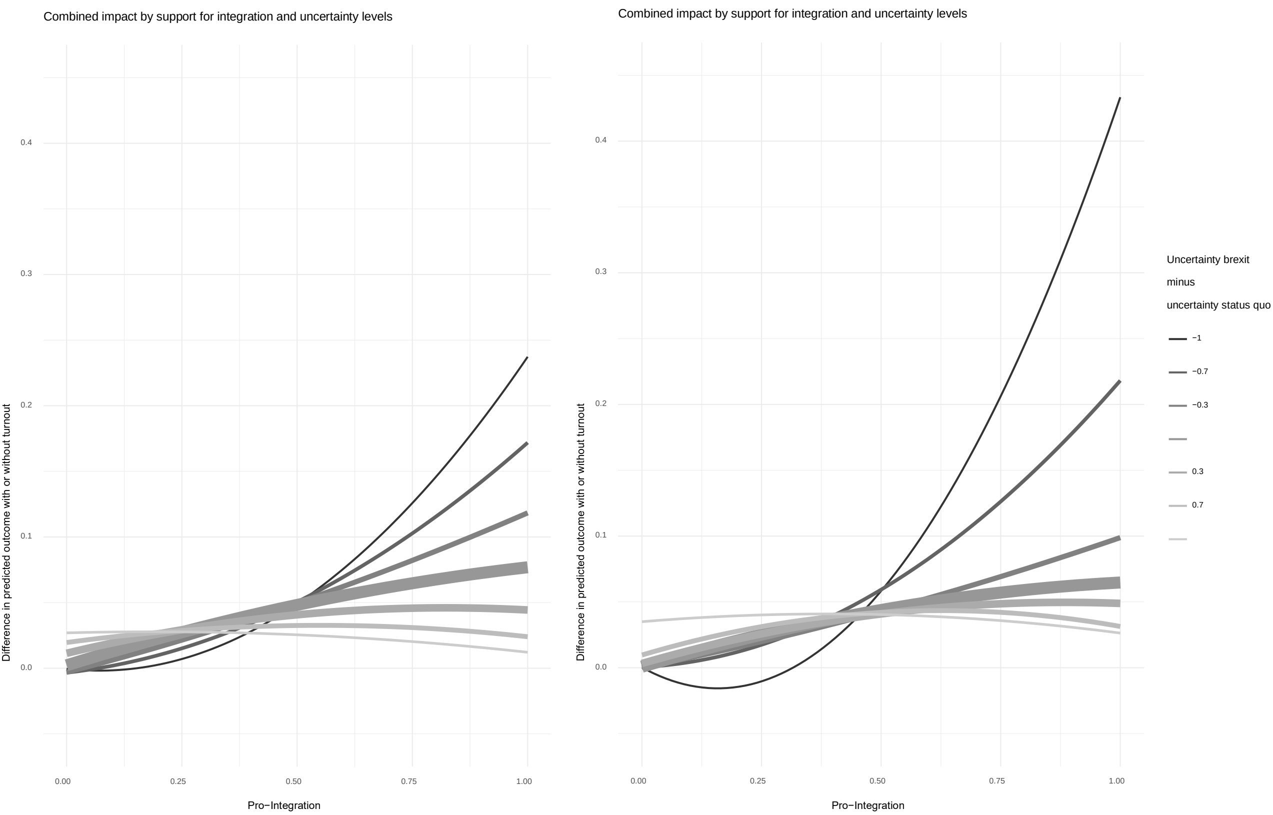

Figure 3: Overall overestimation in impact when ignoring turnout for the 'Brexit' referendum, using smoothened curves. The difference is calculated as the probability of a 'remain' vote calculated for voters only, minus the probability of a 'remain' vote multiplied by the probability to turn out. On the left predicted vote for 'remain' based on median values on all remaining variables, on the right based on the empirical distribution of all variables in the sample. Thickness of lines is proportional to the frequency in the sample.

VI. Conclusion

We present a new theoretical model describing how uncertainty in a referendum impacts twice, affecting turnout and vote choice, and in a direction that exacerbates the impact on the overall outcome of the election. The model predicts that voters who generally prefer one of the two options, but who are positioned far away from both options, tend to abstain when uncertainty about their preferred option is large. Greater uncertainty about its consequences not only reduces the value of an option, but the difference in value between the two options also vanishes for more 'distant' voters. Uncertainty, thus, has a double effect: potential supporters of one option are less likely to vote; and citizens who vote are less likely to support that option.

Our theoretical model assumes risk averse voters and avoids assumptions more common in existing theoretical models. Most current models assume that the voter calculates the probability of being pivotal—a probability that is in fact always negligibly small. Furthermore, they assume an optimal outcome for all voters and a rational calculation by voters to obtain the optimal outcome. Our model is therefore a significant contribution to the theoretical literature on turnout and vote choice in binary elections and referendums.

We evaluate our theoretical model using a survey after the 'Brexit' referendum in the UK. The empirical analysis not only confirms our theoretical expectations, it also provides new insights into voting behavior in the referendum. We find that turnout among 'Brexit' voters would have been higher and even greater numbers would have voted for 'Brexit', if uncertainty around the 'Brexit' option had been smaller. Our analysis also demonstrates the importance of taking turnout into account when explaining or predicting voting behavior. In the 'Brexit' referendum, ignoring turnout leads to an overestimation of support for the status quo by 2.7% on average across our sample—and much higher for particular groups of voters.

We apply the model to the 'Brexit' referendum, but the theory is general. It can be applied to any referendum or election with two candidates or two parties. In fact, the original inspiration for this paper were the referendums on the Nice Treaty in Ireland in 2001–02, where a change in turnout resulted in a radically different outcome when the same referendum was repeated a year later. We thus continue the endeavour by Sanders (1998) and Sattler and Urpelainen (2012) to encourage empirical researchers to combine turnout and vote choice in single model specifications. Extensions to multi-party settings are left for future research.

The findings do not only affect how we should model voting behavior in referendums and elections with two candidates, but it also has implications for democracy and legitimacy more generally. The magnified role of uncertainty in referendums suggests that creating confusion and providing misinformation—or "fake news"—can be a successful strategy in referendum or election campaigns, which undermines the deliberative quality of the democratic process. Secondly, the mechanism as outlined in our model leads to a misalignment between the referendum outcome and the preference distribution of the electorate, whereby a minority potentially manages to impose its preferred option although a majority would be better off with the losing alternative.

APPENDIX

Proof of the Proposition

The set V contains all i such that

$$(\rho_H - \rho_L)[i(\Delta_1 - \Delta_0) - t(\Delta_1 - \alpha\Delta_0)] \geq 0.$$

In equilibrium, the set of voters V and the beliefs ρ_H and ρ_L must be consistent. Note that we always have $i_L \leq i_H$.[19] This means that, conditional on having all citizens participating, we have $(\rho_H - \rho_L) < 0$.

Consider any $\Delta_1 \in [\alpha\Delta_0, \Delta_0]$. The only equilibrium, in this case, is such that every citizen participates and $(\rho_H - \rho_L) < 0$. Indeed, under such circumstances, (12) is satisfied for all i and the beliefs are consistent with the fact that $i_L \leq i_H$ so that more voters support t_1 when $s = L$.

Consider any $\Delta_1 < \alpha\Delta_0$. Assume that the voters' beliefs are such that $(\rho_H - \rho_L) > 0$. Then, and according to (12), the voters who may not participate are located on the extreme right. This, together with the fact that $i_L \leq i_H$, would mean that option t_1 is less likely to be selected when $s = H$, which is not consistent with the initial beliefs. Assume now that the voters' beliefs are such that $(\rho_H - \rho_L) < 0$. According to (12), voters located on the extreme left may not participate. This, in turn, is consistent with the fact that option t_1 is less likely to be selected when $s = H$.

Consider finally any $\Delta_1 > \Delta_0$. Assume that the voters' beliefs are such that $(\rho_H - \rho_L) > 0$. According to (12), voters who may not participate are located on the extreme left. This, together with the fact that $i_L \leq i_H$, would mean that option t_1 is less likely to be selected when $s = H$ which is not consistent with the initial beliefs. Assume now that the voters' beliefs are such that $(\rho_H - \rho_L) < 0$. According to (12), voters located on the extreme right may not participate.[20] This, in turn, is consistent with the fact that option t_1 is less likely to be selected when $s = H$.

---

[19] Indeed, this inequality holds provided $[(t - \Delta_1) - (\alpha t - \Delta_0)][(t + \Delta_1) - (\alpha t + \Delta_0)] \geq 0$, which is always true.

[20] To be precise, notice that when $\Delta_0 < \Delta_1 < \Delta_0 \frac{1-\alpha t}{1-t}$ everyone participates. But when $\Delta_1 > \Delta_0 \frac{1-\alpha t}{1-t}$ those who do not participate are located at the extreme right.

Derivation of the empirical model

To derive the latent regression (14), we turn equation (13) into a latent regression by adding an error term, ϵ_1, to capture measurement error and exogenous factors that are not part of our theoretical model, leading to

$$i(\Delta_0 - \Delta_1) - E(t_0)\Delta_0 + E(t_1)\Delta_1 + \epsilon_1 \geq 0,$$

where $E(t_0) = \alpha t$ and $E(t_1) = t$. By adding coefficients, the constitutive term i and control variables to the regression equation, we get the specification as expressed in equation (14) in the text.

To derive the latent regression (15), we turn equation (6) into a latent regression by again adding an error term, ϵ_2, leading to

$$\delta + (\Delta_0^2 - \Delta_1^2) + 2i\big(E(t_1) - E(t_0)\big) + \epsilon_2 \geq 0,$$

where $E(t_0) = \alpha t$ and $E(t_1) = t$ and $\delta = -t^2(1 - \alpha^2)$. By adding coefficients, the constitutive terms i and $\big(E(t_1) - E(t_0)\big)$, and control variables to the regression equation, we obtain the specification as expressed in equation (15) in the text.

Full regression model

	(1)		(2)		(3)	
	turnout	remain	turnout	remain	turnout	remain
Pro integration (*I*)	−0.459***	3.411***	−0.339***	3.445***	−0.442***	3.763***
	(0.053)	(0.133)	(0.075)	(0.167)	(0.086)	(0.173)
Uncertainty Remain (Δ_1)	−0.193*		−0.204		−0.067	
	(0.106)		(0.158)		(0.187)	
Uncertainty Leave (Δ_0)	−0.379***		−0.244		−0.236	
	(0.108)		(0.162)		(0.194)	
$I \times (\Delta_0 - \Delta_1)$	0.542***		0.703**		0.652**	
	(0.179)		(0.270)		(0.313)	
Impact Brexit ($t_1 - t_0$)		1.338***		1.397***		1.442***
		(0.076)		(0.110)		(0.109)
$(t_1 - t_0) \times I$		1.597***		1.415***		1.181***
		(0.169)		(0.220)		(0.095)
$(\Delta_0^2 - \Delta_1^2)$		1.287***		1.216***		1.050***
		(0.058)		(0.081)		(0.237)
Political knowledge					0.012	0.049***
					(0.015)	(0.012)
Efficacy					0.190***	
					(0.020)	
Education			0.056	0.289***	−0.017	0.319***
			(0.048)	(0.036)	(0.059)	(0.041)
Income			0.036***	0.031***	0.021***	0.026***
			(0.006)	(0.006)	(0.007)	(0.006)
Female			0.055	0.025	0.103**	0.070*
			(0.043)	(0.034)	(0.051)	(0.040)
Age			0.006	0.002	0.017*	−0.0005
			(0.008)	(0.002)	(0.009)	(0.002)
Age squared			0.0001		−0.00002	
			(0.000)		(0.0001)	
Intercept	2.144***	−2.223***	1.206***	−2.688***	0.657***	−2.818***
	(0.042)	(0.056)	(0.187)	(0.196)	(0.224)	(0.185)
N	21,607		13,388		10,855	

*p < .1; **p < .05; ***p < .01

Table A1: Heckman selection models for voting behavior in the 'Brexit' referendum. Both the selection mechanism, turnout, and the outcome, voting for remain, are modelled using probit regression specifications. The turnout and vote choice equations are estimated jointly using Maximum Likelihood estimation.

Alternative model specifications

	(1)		(2)		(3)	
	turnout	remain	turnout	remain	turnout	remain
Pro integration (*I*)	−0.431***	−0.325***	−0.325***	1.273***	−0.436***	1.332***
	(0.053)	(0.028)	(0.075)	(0.035)	(0.035)	(0.039)
Uncertainty Remain (Δ_1)	−0.196*		−0.214		−0.077***	
	(0.106)		(0.157)		(0.188)	
Uncertainty Leave (Δ_0)	−0.403***		−0.260		−0.254	
	(0.108)		(0.162)		(0.195)	
$I \times (\Delta_0 - \Delta_1)$	0.587***		0.733***		0.676**	
	(0.177)		(0.267)		(0.313)	
Impact Brexit ($t_1 - t_0$)		0.400***		0.410***		0.425***
		(0.014)		(0.018)		(0.019)
$(t_1 - t_0) \times I$		−0.364***		−0.382***		−0.460***
		(0.032)		(0.042)		(0.046)
$(\Delta_0^2 - \Delta_1^2)$		0.223***		0.211***		0.205***
		(0.010)		(0.014)		(0.016)
Political knowledge					0.012	0.010***
					(0.015)	(0.002)
Efficacy					0.188***	
					(0.020)	
Education			0.049	0.060***	–0.018	0.063***
			(0.048)	(0.006)	(0.059)	(0.0007)
Income			0.035***	0.007***	0.021***	0.006***
			(0.006)	(0.001)	(0.007)	(0.001)
Female			0.056	0.003	0.105**	0.013*
			(0.043)	(0.006)	(0.051)	(0.007)
Age			0.008	0.0003	0.017*	–0.0002
			(0.008)	(0.0002)	(0.009)	(0.0003)
Age squared			0.000		–0.00002	
			(0.000)		(0.0001)	
Intercept	2.140***	−0.104***	1.176***	–0.193***	0.668***	–0.208***
	(0.042)	(0.012)	(0.186)	(0.025)	(0.224)	(0.030)
N	21,607		13,388		10,855	

*p < .1; **p < .05; ***p < .01

Table A2: Heckman selection models for voting behavior in the 'Brexit' referendum. The selection mechanism, turnout, is modelled using a probit regression; the outcome mechanism, voting for remain, as a linear probability model. The turnout and vote choice equations are estimated jointly using Maximum Likelihood estimation.

	(1)		(2)		(3)	
	turnout	remain	turnout	remain	turnout	remain
Pro integration (*I*)	-0.561^{***}	1.181^{***}	-0.421^{***}	1.207^{***}	-0.436^{***}	1.332^{***}
	(0.018)	(0.035)	(0.025)	(0.045)	(0.086)	(0.039)
Uncertainty Remain (Δ_1)	-0.251^{***}		-0.296^{***}		−0.077	
	(0.033)		(0.050)		(0.188)	
Uncertainty Leave (Δ_0)	-0.435^{***}		-0.285^{***}		−0.254	
	(0.034)		(0.024)		(0.195)	
$I \times (\Delta_0 - \Delta_1)$	0.658^{***}		0.830^{***}		0.676^{**}	
	(0.057)		(0.087)		(0.313)	
Impact Brexit ($t_1 - t_0$)		0.385^{***}		0.399^{***}		0.425^{***}
		(0.019)		(0.024)		(0.019)
$(t_1 - t_0) \times I$		-0.276^{***}		-0.325^{***}		-0.460^{***}
		(0.040)		(0.054)		(0.046)
$(\Delta_0^2 - \Delta_1^2)$		0.229^{***}		0.228^{***}		0.205^{***}
		(0.014)		(0.018)		(0.016)
Political knowledge					0.012	0.010^{***}
					(0.015)	(0.002)
Efficacy					0.188^{***}	
					(0.020)	
Education			0.057^{***}	0.064^{***}	−0.018	0.063^{***}
			(0.016)	(0.008)	(0.059)	(0.007)
Income			0.044^{***}	0.010^{***}	0.021^{***}	0.006^{***}
			(0.002)	(0.001)	(0.007)	(0.001)
Female			0.074^{***}	0.006	0.105^{**}	0.013^{*}
			(0.014)	(0.008)	(0.051)	(0.007)
Age			0.010^{***}	0.002^{***}	0.017^{*}	−0.0002
			(0.003)	(0.0002)	(0.009)	(0.0003)
Age squared			0.000^{***}		−0.00002	
			(0.000)		(0.0001)	
Intercept	1.269^{***}	-0.144^{***}	0.066^{***}	-0.313^{***}	0.668^{***}	-0.208^{***}
	(0.013)	(0.016)	(0.064)	(0.033)	(0.224)	(0.030)
N	21,607		13,388		10,855	

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table A3: Heckman selection models for voting behavior in the 'Brexit' referendum. The selection mechanism, turnout, is modelled using a probit regression; the outcome mechanism, voting for remain, as a linear probability model. Observations are weighted by turnout, to correct for the oversampling of voters. The turnout and vote choice equations are estimated jointly using Maximum Likelihood estimation.

REFERENCES

Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters 80*(1), 123–129.

Arnorsson, A. and G. Zoega (2018). On the causes of brexit. *European Journal of Political Economy*. In press.

Battaglini, M., R. B. Morton, and T. R. Palfrey (2008). Information aggregation and strategic abstention in large laboratory elections. *American Economic Review 98*(2), 194–200.

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

Berry, W. D., J. H. DeMeritt, and J. Esarey (2010). Testing for interaction in binary logit and probit models: Is a product term essential? *American Journal of Political Science 54*(1), 248–266.

Bessi, A., F. Zollo, M. Del Vicario, A. Scala, G. Caldarelli, and W. Quattrociocchi (2015). Trend of narratives in the age of misinformation. *PloS one 10*(8), e0134641.

Boomgaarden, H. G., A. R. Schuck, M. Elenbaas, and C. H. De Vreese (2011). Mapping EU attitudes: Conceptual and empirical dimensions of euroscepticism and EU support. *European Union Politics 12*(2), 241–266.

Bursztyn, L., D. Cantoni, P. Funk, and N. Yuchtman (2018). Polls, the press, and political participation: The effects of anticipated election closeness on voter turnout. Technical report, National Bureau of Economic Research. Working Paper 23490.

Campbell, A., P. E. Converse, W. E. Miller, and D. E. Stokes (1960). *The American Voter*. Chicago: Wiley.

Carpini, M. X. D. and S. Keeter (1996). *What Americans Know about Politics and Why it Matters*. New Haven: Yale University Press.

Clarke, H. D., M. Goodwin, and P. Whiteley (2017). Why britain voted for brexit: An individuallevel analysis of the 2016 referendum vote. *Parliamentary Affairs 70*(3), 439–464.

Colantone, I. and P. Stanig (2018). Global competition and brexit. *American Political Science Review 112*(2), 201–218.

de Vreese, C. H., R. Azrout, and H. G. Boomgaarden (2019). One size fits all? testing the dimensional structure of eu attitudes in 21 countries. *International Journal of Public Opinion Research 31*(2), 195–219.

Degan, A. and A. Merlo (2011). A structural model of turnout and voting in multiple elections. *Journal of the European Economic Association 9*(2), 209–245.

Del Vicario, M., A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences 113*(3), 554–559.

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

Denny, K. and O. Doyle (2008). Political interest, cognitive ability and personality: Determinants of voter turnout in britain. *British Journal of Political Science 38*(2), 291–310.

Dimitrova, D. V., A. Shehata, J. Strömbäck, and L. W. Nord (2014). The effects of digital media on political knowledge and participation in election campaigns: Evidence from panel data. *Communication Research 41*(1), 95–118.

Dür, A. and N. Konstantinidis (2013). International treaty ratification and party competition: Theory and evidence from the eu's constitutional treaty. *Political Science Research and Methods 1*(2), 179–200.

Elkink, J. A. and R. Sinnott (2015). Political knowledge and campaign effects in the Irish vote on the Lisbon Treaty. *Electoral Studies 38*, 217–225.

Farrell, D. M. and R. Schmitt-Beck (Eds.) (2003). *Do Political Campaigns Matter? Campaign effects in elections and referendums*. London: Routledge.

Feddersen, T. J. and W. Pesendorfer (1996). The swing voter's curse. *American Economic Review 86*(3), 408–424.

Feddersen, T. J. and W. Pesendorfer (1999). Abstention in elections with asymmetric information and diverse preferences. *American Political Science Review 93*(2), 381–398.

Fieldhouse, E., J. Green, G. Evans, H. Schmitt, C. van der Eijk, J. Mellon, and C. Prosser (2015). British election study internet panel wave 10.

Gabel, M. J. (1998). Public support for European integration: An empirical test of five theories. *Journal of Politics 60*(2), 333–354.

Galam, S. (2017). Marine le pen can breach her glass ceiling: The drastic effect of differentiated abstention. arXiv:1703.04643.
Ghirardato, P. and J. Katz (2002). Indecision theory: Quality of information and voting behavior. https://dx.doi.org/10.2139/ssrn.260580.
Goodwin, M. and C. Milazzo (2017). Taking back control? investigating the role of immigration in the 2016 vote for brexit. *British Journal of Politics and International Relations 19*(3), 450–464.
Goodwin, M. J. and O. Heath (2016). The 2016 referendum, brexit and the left behind: An aggregate-level analysis of the result. *Political Quarterly 87*(3), 323–332.

Grant, D. and M. Toma (2008). Elemental tests of the traditional rational voting model. *Public Choice 137*(1-2), 173–195.
Hobolt, S. B. (2005). When europe matters: The impact of political information on voting behaviour in eu referendums. *Journal of Elections, Public Opinion and Parties 15*(1), 85–109.

Hobolt, S. B. (2009). *Europe in Question: Referendums on European integration*. New York: Oxford University Press.

Hobolt, S. B. (2016). The brexit vote: A divided nation, a divided continent. *Journal of European Public Policy 23*(9), 1259–1277.
Hug, S. (2002). *Voice of Europe: Citizens, referendums, and European integration*. Rowman and Littlefield.

Hug, S. and T. Schulz (2007). Referendums in the eu's constitution building process. *Review of International Organizations 2*(2), 177–218.
Kenski, K. and N. J. Stroud (2006). Connections between internet use and political efficacy, knowledge, and participation. *Journal of Broadcasting & Electronic Media 50*(2), 173–192.
Kim, J. and M. Fey (2007). The swing voter's curse with adversarial preferences. *Journal of Economic Theory 135*(1), 236–252.
Larcinese, V. (2009). Information acquisition, ideology and turnout: theory and evidence from britain. *Journal of Theoretical Politics 21*(2), 237–276.
Lau, R. R. and D. P. Redlawsk (1997). Voting correctly. *American Political Science Review 91*(3), 585–598.
Lau, R. R. and D. P. Redlawsk (2001). Advantages and disadvantages of cognitive heuristics in political decision making. *American Journal of Political Science*, 951–971.
Levine, D. K. and A. Mattozzi (2018). Voter turnout with peer punishment. Working Paper, European University Institute.

Matsusaka, J. G. (1995). Explaining voter turnout patterns: An information theory. *Public Choice 84*(1/2), 91–117.
Matti, J. and Y. Zhou (2017). The political economy of brexit: Explaining the vote. *Applied Economics Letters 24*(16), 1131–1134.
Myatt, D. P. (2015). A theory of voter turnout. Working Paper, London Business School.

Nicholson, S. P. (2003). The political environment and ballot proposition awareness. *American Journal of Political Science 47*(3), 403–410.
Nie, N. H., J. Junn, and K. Stehlik-Barry (1996). *Education and Democratic Citizenship in America*. Chicago: University of Chicago Press.

Nie, N. H., S. Verba, and J. R. Petrocik (1979). *The Changing American Voter*. Cambridge: Harvard University Press.

Oliveros, S. (2013). Abstention, ideology and information acquisition. *Journal of Economic Theory 148*(3), 871–902.
Reif, K. and H. Schmitt (1980). Nine second-order national elections: A conceptual framework for the analysis of european election results. *European Journal of Political Research 8*(1), 3–44.

Sanders, M. S. (1998). Unified models of turnout and vote choice for two-candidate and three-candidate elections. *Political Analysis 7*, 89–115.
Sattler, T. and J. Urpelainen (2012). Explaining public support for international integration: How do national conditions and treaty characteristics interact with individual beliefs? *Journal of Politics 74*(4), 1108–1124.

Schneider, G. and L.-E. Cederman (1994). The change of tide in political cooperation: A limited information model of european integration. *International Organization 48*(4), 633–662.
Schneider, G. and P. A. Weitsman (1996, 4). The punishment trap: Integration referendums as popularity contests. *Comparative Political Studies 28*(54), 586–607.
Signorino, C. (1999). Strategic interaction and the statistical analysis of international conflict. *American Political Science Review 93*, 279–298.
Sinnott, R. (2003). Attitudes and behaviour of the irish electorate in the second referendum on the treaty of nice. Report for the European Commission Representation in Ireland, University College Dublin.

Smets, K. and C. van Ham (2015). The embarrassment of riches? a meta-analysis of individuallevel research on voter turnout. *Electoral Studies 32*(2), 344–359.

Solon, G., S. J. Haider, and J. M. Wooldridge (2015). What are we weighting for? *Journal of Human Resources 50*(2), 301–316.

Toomet, O. and A. Henningsen (2008). Sample selection models in R: Package sampleSelection. *Journal of Statistical Software 27*(7).

Tucker, J. A., A. C. Pacek, and A. J. Berinsky (2002). Transitional winners and losers: Attitudes toward eu membership in post-communist countries. *American Journal of Political Science*, 557–571.

Vasilopoulou, S. (2016). Uk euroscepticism and the brexit referendum. *Political Quarterly 87*(2), 219–227.

Walter, S., E. Dinas, I. Jurado, and N. Konstantinidis (2018). Noncooperation by popular vote: Expectations, foreign intervention, and the vote in the 2015 greek bailout referendum. *International Organization 72*(4), 969–994.

Wickham, H. (2015). *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Zaller, J. (1992). *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.

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