

# Forecasting Candidate States' Compliance with EU Accession Rules, 2017-50

Tobias Böhmelt

*ETH Zürich (CH) and University of Essex (UK)*

Department of Government

Wivenhoe Park, Colchester

United Kingdom

[tobias.boehmelt@ir.gess.ethz.ch](mailto:tobias.boehmelt@ir.gess.ethz.ch)

[tbohmelt@essex.ac.uk](mailto:tbohmelt@essex.ac.uk)

Tina Freyburg

*University of St. Gallen (CH)*

School of Economics and Political Science

Müller-Friedbergstrasse 8, St. Gallen

Switzerland

[tina.freyburg@unisg.ch](mailto:tina.freyburg@unisg.ch)

**Acknowledgments:** We thank Lena Kiesewetter, Chingun Anderson, and Nicolas Cerkez for their research assistance. We are also grateful to Solveig Richter and Frank Schimmelfennig for feedback, while the journal's editors and the anonymous reviewers provided valuable comments that helped us to improve the article.

The EU is said to be tired of enlargement – but how likely is it that a candidate would be ready to join within the next ten, fifteen, or more years? This research forecasts how prospective members are likely able to perform in implementing EU law until 2050. Using data on the alignment patterns of all states from the 2004, 2007, and 2013 accession rounds as well as of five current and potential candidate countries, we construct an empirical model on candidates’ ability to comply with the *acquis communautaire*. We employ in-sample and out-of-sample techniques to assess the model’s prediction accuracy. After demonstrating that this accuracy is high and that we improve upon a naïve approach, we ultimately forecast the five candidates’ potential compliance levels in 2017-50. Our research shows that *only one* of these states might sufficiently be able to comply with the accession criteria until 2023, while many of them are unlikely to be ready before the mid-2030s. Due to the focus on prediction and forecasting, our contribution is given by the research’s policy relevance and its methodological innovation.

**Keywords:** *acquis communautaire*, accession conditionality, European Union, forecasting, prediction

## **Introduction**

‘Enlarging the EU has taken a back seat as members grapple with problems closer to home.’<sup>1</sup> While this may seem like a contemporary statement about European Union (EU) enlargement in times of crisis, it is nonetheless how the editor of the Financial Times, Lionel Barber, summarized the situation in the mid-1990s. Although the EU was generally committed to opening its gates to the East back then, it was careful not to encourage the central and east European countries to push for membership (see also Mattli and Plümper 2002). Indeed, Barber goes on by quoting a senior Commission official saying the ‘[EU’s current] level of seriousness about enlargement is not minimal, it simply does not exist.’

This episode illustrates that enlargement historically stems from the pressure of countries that aspire to join, not from an expansionist ambition on the EU’s side. Moreover, it emphasizes that the desirability of enlargement and its potentially destabilizing effects have been subject to heated debates since the earliest days of the European Communities (O’Brennan 2014, 223f). And yet, parallel to its institutional deepening, the EU widened significantly from originally six members to 28 by 2017. This raises key analytical and empirical questions that this article seeks to address: what would be an adequate baseline model for predicting candidate countries’ ability to comply with EU law? And, against previous and on-going enlargement experiences, how likely is it that a candidate would be ready or able to join over the next decade(s)?

EU enlargement policy is path-dependent due to the self-reinforcing nature of a series of sequenced decisions and, hence, rather difficult to reverse (Vachudova 2005; Giandomenico 2009; Szymański 2013). The institutional architecture of EU

---

<sup>1</sup> ‘Brussels keeps shut the gates to the East’ (Lionel Barber), Financial Times (November 16, 1995), 17.

enlargement policy with its hybrid nature adds to this rigidity. Each member state has, in theory, the possibility to veto at multiple stages. Yet, enlargement cannot take place without the approval of central EU institutions, notably the European Parliament and the Commission. These supranational institutions are strongly committed to ‘the firm prospect of EU membership’ (European Commission 2016). Moreover, the standardized procedure also obliges them to consider all applications from European countries according to the same standards. This rigid architectural design of EU enlargement policy allows us to presume a certain continuity against which we can assess a candidate country’s ability to comply. That is, if EU enlargement policy is generally consistently linked to compliance in the target countries (Szymański 2013; Schimmelfennig 2008) and it is, then and primarily, properties of candidate countries that explain variance in compliance ability, models on candidates’ past compliance can, in principle, accurately predict and forecast their future ability to comply.<sup>2</sup>

Drawing on previous statistical accounts of candidate countries’ compliance with EU accession requirements (notably, Hille and Knill 2006; Schimmelfennig and Scholtz 2008; Toshkov 2008; Levitz and Pop-Eleches 2010; Böhmelt and Freyburg 2013; 2015), we consider a set of core exogenous predictors to construct a baseline model for the forecast. We examine the predictive power of this model via in-sample and out-of-sample techniques, including a four-fold cross-validation exercise. Finally, after having demonstrated that prediction accuracy is given and that we are able to improve upon a naïve approach, we forecast five current and potential EU candidate countries’ (Albania, Bosnia-Herzegovina, the Former Yugoslav Republic of (FYR) Macedonia, Serbia, and Turkey) compliance levels with the EU *acquis communautaire* until 2050.

---

<sup>2</sup> Note already here that this research is about states’ ability to comply or adjust, not the political willingness to actually do so.

Over the past few decades, the literature has developed influential theoretical frameworks that shed light on the underlying mechanisms of candidate states' compliance with the EU *acquis*, and has empirically tested these frameworks. However, the empirical evidence is often conducted *ex-post* on observed data. Despite important insights, there are key limitations associated with these kinds of empirical studies. As Ward et al. (2010, 364) forcefully remind us, policy prescriptions cannot be 'based on statistical summaries of probabilistic models.' Hence, drawing inferences based on statistical significance testing alone might be misleading and not inform us about the (predictive) power of an empirical model (see also Hegre et al. 2017). While statistically significant results may improve our understanding of the relationship between variables in each sample under study, they could not provide information on the exact same relationship in another, i.e., new sample of data – like the future. Prediction and forecasting methods can help address these shortcomings. In the words of Schneider et al. (2010, 1): 'anticipating the future is both a social obligation and intellectual challenge that no scientific discipline can escape,' while predictive power is usually regarded as the 'gold standard' for assessing the quality of explanatory models in most disciplines, including the social sciences (Goertz 2006).

Following Hegre et al. (2017, 114), we define 'forecasts as predictions about unrealized outcomes given model estimates from realized data. [...] 'Prediction' is a more general concept, and refers to the assignment of a probability distribution to an outcome based on such model estimates, but may be applied to realized as well as unrealized outcomes. More colloquially, forecasts are predictions about tomorrow given information we have about what has happened up to today. This means two inputs are required to make forecasts: realized data and estimators; and one output is produced: predictions.' Against this background, we consider prediction and

forecasting techniques not only as scientifically valuable, but also as practically highly useful due to their value.<sup>3</sup>

Overall, we therefore seek to contribute to both the academic and the policy-oriented literature. First, our work provides an assessment when current and potential candidates might be ready for EU membership in light of their ability to adjust to the *acquis*, if at all, through systematic research based on information on the previous, current, and potential candidates' levels of compliance. Since earlier work has paid little attention to predicting and forecasting states' ability to implement EU law, policymakers lack guidance for assessing the success of EU enlargement politics and, hence, making an informed statement on potential readiness of candidates for future accessions.

Second, we develop a model that enables forecasts of candidate countries' compliance levels, i.e., their ability to comply with the *acquis* that future research can easily rely on, verify, and extend. Specifically, we predict candidate states' compliance levels in the future after having determined that our set of exogenous factors (including fixed effects) predicts accurately observed compliance patterns with genuine in-sample and out-of-sample techniques. However, neither is the predictive model in  $t$  based on data in  $t$  nor do we use earlier compliance patterns to forecast compliance in the future (for a discussion of the importance on this, see, e.g., Desmarais and Cranmer 2013). To this end, we provide a comprehensive discussion

---

<sup>3</sup> Having said that, it should be kept in mind that prediction differs from explanation. In the words of Hegre et al. (2017, 115): '[w]hen evaluating the relationship between prediction and explanation it is important to recognize the different purposes of forecasting. Forecasting can help researchers to test, improve, and build their theories. However, forecasting not only fulfils scientific objectives; it also enables policymakers to formulate evidence-based policies regarding peace and security issues. Forecasts can help designing polices or act merely as an early-warning tool.'

of the conceptualization of compliance used, the specifications of the explanatory variables identified in earlier work, and, most crucially, previous findings. As we provide clear guidelines for prediction exercises in general, we also seek to contribute to the ‘forecasting literature’ in EU politics (Thomson and Hosli 2006; Bechtel and Leuffen 2010) and to enrich the debate on the validity of policies based on empirical models (Goldstone et al. 2010; Schneider et al. 2011; Ward et al. 2010).

Using the 2004 accession round as a benchmark, the results show that *only one country* of the current and potential EU candidates is likely to be able to sufficiently comply with the accession criteria until 2023, FYR Macedonia, while most of them may not be ready for accession before the mid-2030s. Our forecast emphasizes that Albania or Bosnia-Herzegovina may even face difficulties in satisfactorily incorporating EU legislation before 2050. Let us emphasize again, however, that this forecast captures the ability of states to comply with the EU *acquis*. We do not forecast actual accession dates or states’ willingness – or unwillingness – to comply with EU law. In light of this, our results might underline that populist foreign policy positions such as the British Vote Leave campaign’s claim that most of the candidate countries would join by 2020 seem mistaken.<sup>4</sup>

### **Path Dependency in EU Enlargement and Candidates’ Compliance with EU Law**

Enlargement policy has been frequently portrayed as a policy that, once established, is increasingly difficult to reverse due to self-reinforcing lock-ins caused by sequenced decision events that make change costly and normatively inappropriate (see, e.g., Vachudova 2005; Giandomenico 2009; Schimmelfennig 2009; Szymański 2013). As a result, member states have *de facto* little option, but to continue the set path despite

---

<sup>4</sup> See <http://www.bbc.com/news/uk-politics-eu-referendum-36355138> (April 10, 2017).

potentially deviating preferences. In fact, the fundament of EU enlargement policy remains mainly unchanged since the 1990s. Modifications so far typically concerned detailed issues, such as the precise requirements for candidates, and were largely motivated by the specificities of the respective candidates rather than EU-internal politics (Szymański 2013; Schimmelfennig 2008). Moreover, it appears that as the process unfolds, and a country is ‘administratively put under the responsibility of DG Enlargement’ (Giandomenico 2009, 111), an ultimate rejection of accession becomes increasingly unlikely (see Böhmelt and Freyburg 2013, 267). Hence, amidst economic and political crises, enlargement policy appears to be ‘running on autopilot.’ Year after year, ‘EU member states continue to make choices [...] that keep the enlargement process going’ (Vachudova 2014, 123). In short, considering the remarkable continuity of EU enlargement policy and its reactive nature, ‘revolutionary changes’ are unlikely to happen.

Of course, the idea of path dependency does not preclude the possibility of institutional or policy changes in the unlikely event of random and unexpected events (‘exogenous shocks’), provided they are sequenced correctly (Thelen and Steinmo 1992). However, anecdotal evidence suggests that even contemporary events, which can be characterized as ‘no equal’ in EU history otherwise, might have no such effect – as illustrated by the following recent enlargement steps: in July 2013, Croatia became a member of the Union; in January 2014, Serbia’s accession negotiations formally began; in February 2016, Bosnia-Herzegovina submitted its application to join the EU; and at the end of June 2016, negotiations on a new chapter were opened with Turkey (#33: financial and budgetary provisions). Overall, it seems reasonable to assume that future enlargement policy will continue to develop at roughly the same historical pace. Against this background, what would be an adequate baseline model



for explaining and, as it is primarily the interest of this article, *predicting* candidate countries' ability to comply?

The EU's pre-accession policy is a demanding exercise for any candidate. As determined by the Copenhagen European Council in 1993, states aspiring to join the EU must bring their laws, regulatory frameworks, and administrative practices in line with the *acquis* prior to accession. Given the amount of EU laws and regulations,<sup>5</sup> as well as the corresponding administrative and institutional requirements, this adjustment process is not simple and the ultimate accession to the EU is uncertain. However, responding to a candidate's application with the opening of accession negotiations, the EU creates an expectation that 'the applicant country will, at some yet unknown moment join' (Steunenberg and Dimitrova 2007, 3), namely once it complies with the EU's rules and regulations. Under what conditions are candidate states likely to be able to formally adopt the EU rules and regulations and apply them in practice?

The existing quantitative research<sup>6</sup> on candidate countries' compliance with EU accession requirements identified a set of variables that vary not only regarding the conceptualization and operationalization of the dependent variable, but also with the list and specifications of the explanatory variables included and, most importantly, the findings. This work can be grouped into studies centring on the EU's effect on the

---

<sup>5</sup> Toshkov, for example, refers to 'more than 100,000 legislative acts' that have to be incorporated into national legal systems, see <http://www.dimiter.eu/Eurlex.html> (April 10, 2017).

<sup>6</sup> By focusing on quantitative research only, our research does not wish to imply that work based on other methods is negligible. Rather, our literature review revealed that existing quantitative research explicitly builds on the key (and statistically testable) insights of the various small/medium-*N* case studies that continue to dominate the field. While the quantitative study of compliance, transposition, and implementation of EU law on the side of member states has flourished (see Toshkov 2010), there still are only a few systematic quantitative studies on compliance of candidate states in the context of EU enlargement.

quality of democracy or democratic governance in candidate states (Schimmelfennig and Scholtz 2008; Levitz and Pop-Eleches 2010) and those seeking to explain variation in the extent to which these countries implemented the *acquis* prior to accession (Hille and Knill 2006; Toshkov 2008; Böhmelt and Freyburg 2013; 2015). We focus on the latter strand, especially the work by Hille and Knill (2006) as well as Böhmelt and Freyburg (2013; 2015) since they share a similar research interest and measure compliance the same way as we do (discussed below). Adopting a rationalist perspective, complemented with managerialists' insights, these studies concentrate on a country's relevant domestic actors' ability, and also partly their willingness, to adopt EU rules and adapt the corresponding national legislation.

Despite the substantive similarities, however, existing studies vary regarding research designs and findings. The empirical analyses differ in country years, with Hille and Knill (2006) studying 13 candidates from 1999 to 2003 and Böhmelt and Freyburg (2013; 2015) covering an extended sample of 16 states in 1998-2009. Furthermore, Böhmelt and Freyburg (2013) employ generalized additive models, while a linear model is used in either Hille and Knill (2006) or Böhmelt and Freyburg (2015). Most covariates are measured with the same data; yet, while Böhmelt and Freyburg (2013, 262) use the World Bank Development Indicators to capture a bureaucracy's financial capabilities, Hille and Knill (2006, 544) rely on data provided by the CIA World Factbook for government expenditures and use the European Commission regular reports for gross domestic product per capita. It seems that the extent to which the relationship between candidates' levels of compliance and the theoretically derived determinants of effective conditionality is statistically significant in regression models considerably varies, too, as shown in the appendix.<sup>7</sup>

---

<sup>7</sup> The appendix lists the covariates included in the models and indicates whether a significant negative

All three studies confirm that conditionality is the main force driving candidates' compliance with the EU *acquis*. However, Hille and Knill (2006, 549) demonstrate that 'the functioning and the quality of the domestic bureaucracy constitute crucial preconditions.' Böhmelt and Freyburg (2015) add that candidate states may free ride on the compliance efforts of others, and their 2013 study reports that, independently from country and EU-level conditions, the leverage of conditionality varies over the pre-accession process. The question remains whether existing explanatory models are also suited to *predict* compliance ability of candidate states.

As indicated above, Ward et al. (2010), among others, emphasize that results in the form of regression coefficients may not tell us much about how candidates' compliance with EU law will develop in the future. Policy prescriptions cannot be based on statistical summaries of probabilistic models. Hypothesis testing that ignores out-of-sample heuristics faces the inherent risk of fitting to a specific sample's idiosyncrasies rather than identifying stable structural relationships between a dependent variable of interest and its determinants. In fact, if a model *explains* the relationship between, in our case, states' ability to comply with EU law and some explanatory factors well in-sample, we merely assume that it also performs well when presented with new data and trying to predict *out-of-sample*.<sup>8</sup> But if the model only gives a description of this relationship in the original data set without capturing underlying causal relations, the chances to make correct predictions with new data are likely to be undermined (Beck et al. 2000). In the words of Ward et al. (2010, 364), 'it

---

(positive), a non-significant negative (positive), or no relationship (~zero) is reported. We distinguish between 'non-significant relationships' and 'no relationship,' as statistical significance does not equal substantive significance. In small samples, for instance, the lack of statistical significance might obscure a substantively important relationship (see Toskov 2010).

<sup>8</sup> See also Desmarais and Cranmer (2013) on the distinction between *explaining* and *predicting*. As indicated, we are interested in the latter.

is quite possible to focus on statistically significant results that are artefacts in the sense that they do not generalize beyond the specific cases studied. This happens if we focus only on statistically significant relationships and may actually hinder our ability to generalize to out-of-sample situations, such as the future!

The following analysis seeks to address this shortcoming. We first identify the necessary variables for our model and then predict candidates' ability to comply with EU law in-sample and out-of-sample. Moving from empirical analyses based on statistical significance to prediction offers a more solid scientific basis for assessing future levels of compliance with EU accession law, which is not only highly relevant both from a policy and scholastic perspective (Bechtel and Leuffen 2010; Schneider et al. 2010; 2011). It also allows us to analyse whether our selected model based on in-sample prediction can also accurately predict candidates' ability to comply out-of-sample, as theoretically suggested by the continuity argument.

## **Research Design**

### *Empirical Strategy and Dependent Variable*

Our predictions and forecasts are based on an OLS regression model that analyses data on candidate states' compliance with EU law using core exogenous predictors, which are meant to primarily proxy the costs and capabilities arguments emphasized in previous empirical studies. Our sample consists of eighteen (previous, current, and potential) candidates for which reliable data are available, as listed in the appendix. We use the country-year as the unit of analysis, while a state drops out of the sample once the Commission does no longer provide progress reports; this happens either at the time when accession treaties are signed or at the end of the time period covered by

this study (with the latest reports published in November 2016, which are included in our analysis).

To arrive at our forecast of states' compliance ability in 2017-50, we complete the following steps. First, we identify a set of potential predictors that may help us in explaining candidates' ability to comply and, thus, their compliance levels. A crucial requirement for these predictors is not only that they are exogenous to our dependent variable, but also have available high-quality projections until 2050. Clearly, not many variables meet these criteria. We follow Hegre et al. (2013) and likewise focus on a set of socio-demographic variables for which projections until 2050 exist and that can be linked to the quantitative literature on effective EU enlargement policy and, especially, previous claims about costs and capabilities. In addition, we add fixed effects for enlargement rounds and states, which then capture temporal shocks, EU policies that are specific for each enlargement rounds, or unobserved unit-level influences that affect compliance.

Second, for determining the prediction accuracy of this model, we estimate it on a time-series cross-sectional sample in 1998-2008 ( $N=120$  country years), which we then assess with in-sample techniques. Third, we examine the out-of-sample prediction power by employing a four-fold cross-validation exercise and comparing our predictions for 2009 to 2016 (based on the estimates for the 1998-2008 period) with observed values (for which we have data). After having assessed the prediction power of our model, and after having concluded that this power is in fact given, the final forecast for the 2017-50 period is based on a model that uses the entire set of already observed data in 1998-2016. We opted for 2050 as the final year in our forecast, as Hegre et al. (2013), and since we have only 19 years of observed data points. We present results for the aggregated sample as well as individual countries.

Figure 1 shows how we analyse the data and combine them with the statistical model to obtain our forecast of candidate states' compliance.

[Figure 1 here]

To quantify and measure candidates' compliance with EU law, our dependent variable, we employ the updated compliance-level data from Böhmelt and Freyburg (2013; 2015) who use the EU Commission's annual reports<sup>9</sup> on each candidate's progress in aligning policies toward EU requirements. In these reports, the Commission explains and assesses in detail what each candidate has achieved over the last year, and identifies areas where more effort is needed to have the 'ability to assume the obligations of membership.' To this end, the *acquis* is divided into 30-40 (depending on the accession round) subject-related chapters. In addition to these policy sectors, each report offers a general evaluation for each country. The Commission has been publishing these regular reports since 1998 for those countries that pertained to the 2004 and 2007 enlargement rounds, and with varying starting years for the other candidates in the sample.

Reporting on a candidate stops the year the accession treaty was signed for the 2004 accession-round states (2003). For Bulgaria and Romania, progress reports are given until 2005 only, but we cover the last year (2006) before the 2007 accession with a joint report.<sup>10</sup> Similarly, the Commission published a 'comprehensive monitoring report on Croatia's state of preparedness for EU membership' in 2012 (i.e., one year after the accession treaty has been signed), which follows the same

---

<sup>9</sup> See [https://ec.europa.eu/neighbourhood-enlargement/countries/package\\_en](https://ec.europa.eu/neighbourhood-enlargement/countries/package_en) (April 10, 2017).

<sup>10</sup> See [https://ec.europa.eu/neighbourhood-enlargement/sites/near/files/pdf/key\\_documents/2006/sept/report\\_bg\\_ro\\_2006\\_en.pdf](https://ec.europa.eu/neighbourhood-enlargement/sites/near/files/pdf/key_documents/2006/sept/report_bg_ro_2006_en.pdf) (April 10, 2017).

structure as the annual progress reports, and we use this file to code the last year before Croatia's accession to the EU and include this country-year in our data set as well. The Commission reports have the advantage that their data quality is high and that they evaluate both formal and practical compliance with EU law of each candidate state on an annual basis in a standardized and comparable manner (Hille and Knill 2006, 541f).

The final dependent variable, a country's logged degree of compliance with EU law in each policy area, is coded along the ordinal four-value assessment provided by the Commission (Böhmelt and Freyburg 2013; 2015): the value of 0 is assigned when a country does not comply with the *acquis* in a specific issue area; 1 if a country partly complies with EU laws and regulations in a specific issue area, although substantially more efforts are necessary; 2 if a country almost fully complies with the *acquis* in a specific issue area, although more efforts are necessary; and 3 when a country fully complies with EU laws and regulations in a specific issue area. Each sector thus receives a value between 0 and 3, while higher values signify higher compliance with the *acquis communautaire*.

Böhmelt and Freyburg (2013; 2015) then estimate the average degree of (logged) compliance for a country in each year by calculating the mean value across all policy areas plus the general evaluation and taking the natural logarithm. Focusing on the 'more general rather than issue or policy-specific' (Hille and Knill 2006, 535) performance of countries, this strategy ensures that we receive a standardized and, hence, comparable measure for all countries at different enlargement stages and rounds. Finally, note that we do, in fact, include the 2016 Commission reports, which comprise compliance patterns that are likely to be affected by populist backlashes of nationalist and populist Euro-scepticism, including the arguably most extreme change

in candidate countries: the authoritarian turn in Turkey that militates against meeting the EU's political accession conditions.

### *Predictors*

We specify a model with core explanatory variables that fulfil three criteria pivotal for our forecasting aim. First, the chosen variables are exogenous to our 'indirect measure' (Toshkov 2008) of compliance performance with EU accession rules based on Commission reports (or they are time-invariant). Second, they arguably proxy the costs and capabilities arguments emphasized by existing statistical studies. And, finally, good projections until 2050 are available for the time-variant items. The selected variables are a time trend, fixed effects for enlargement rounds, country fixed effects, the demographic composition of a state, infant mortality, and education (Table 1).

[Table 1 here]

Commonly used operationalizations and variable specifications in earlier studies may suffer from possible endogeneity with our outcome variable, candidates' compliances as reported by the Commission, for predominantly two reasons. First, the Commission stresses the need for administrative and judicial capacity to ensure correct implementation and application of the many rules next to the actual adoption of the *acquis* (Christoffersen 2007, 47). Hence, dependent and core explanatory variables in previous work might conceptually overlap. Second, the expert scores used to measure some determinants of candidates' compliance might also inform the Commission's assessment of compliance with its accession rules or have been



informed by its progress reports (Toshkov 2008, 382). For instance, state capacity is frequently measured by the expert ratings collected for the World Bank; political costs or incentives are often operationalized with data on a country's level of democracy from the Polity IV project. The Commission uses many sources including contributions from the respective candidate government, the member states, the EU Parliament reports, as well as information from various international and non-governmental organizations (Christoffersen 2007, 31). At the same time, the Commission makes 'efforts to ensure that international organizations such as IMF and World Bank pay attention to the reports' (Kelley 2006, 34).

Moreover, for commonly used operationalizations, notably the anticipated adoption costs based on a country's level of political and economic liberalization or its capabilities in terms of bureaucratic strength or government expenditures, reliable projections are not available for the period 2017-50. Therefore, we use projections for predictors – demographics, infant mortality, and education – from the UN World Population Prospects and the International Institute for Applied Systems Analysis (IIASA) that we believe can serve as proxies for commonly suggested covariates. The World Population Prospects provides 'estimates of demographic indicators for all states in the international system' (Hegre et al. 2013, 254) and projections for these key variables until 2050. These projections are based on national population censuses and revised in consultation with experts from national projection-making agencies every two years. We take the mean scenarios of the UN projections and complement them with those provided by the Austrian non-governmental research organization IIASA based on expert and argument-based forecasting, in particular its 2001 revisions of the World Population Program, which was released in its final form in

2004.<sup>11</sup> Despite inherent demographic uncertainty, the estimates of demographic indicators provided by both UN and IIASA are seen as the most authoritative (O’Neill et al. 2001, 206; see also Hegre et al. 2013).

Our first cluster of predictors seeks to proxy the domestic costs associated with compliance by a target government. A government’s preferences are likely to be influenced by the extent to which the public supports EU membership and, hence, the likelihood of the public to punish potentially costly reforms in coming elections.<sup>12</sup> Various studies on the relationship between societal characteristics, mass attitudes, and EU integration/membership, respectively, emphasize socio-economic determinants of EU support, with some groups gaining and others losing from membership. From this utilitarian perspective, an individual’s expected net gain from EU membership significantly depends on her individual characteristics, notably *age* or *education*. While studies of attitudes in the then-current EU members consistently demonstrate that ‘winners,’ commonly the young and better educated, are more likely to have favourable EU attitudes (Gabel 1998), the evidence for such individual-level drivers of EU support in candidate countries remains mixed (see also Börzel et al.

---

<sup>11</sup> For more information on global population projections and how they are produced, see the Guide to Global Population Projections by O’Neill et al. (2001). We also performed our calculations using the data on other UN scenarios. Our results based on these calculations do not qualitatively differ from those reported below.

<sup>12</sup> An anonymous reviewer suggested that support for EU membership is only one factor regarding adjustment costs and that the variables we use may in fact be more related to perceived legitimacy; First, while the latter is different from the former, the two concepts are related. Second, strong support might incline a government to accept high adjustment cost, but it does not tell us how high these costs are originally. High support reduces the net costs, but it may not serve as a (comparative) indicator of compliance costs. Note, however, that we focus on prediction rather than explanation and, thus, the specific relationship a single predictor has with the outcome variable, i.e., positive, negative, or none, does ultimately not matter. What matters is whether this predictor, as part of the full model, contributes to the prediction and forecasting power of the model, and we demonstrate this in the following sections.

2010). In fact, studies relying on the Central and Eastern Eurobarometer survey data find only weak and cross-nationally inconsistent effects of demographic characteristics on support for EU membership (Tucker et al. 2002, 569; Tverdova and Anderson 2004). Work using alternative data sources, namely the actual results of the referendums at the regional level, however, find that higher level of education increases the likelihood of voting in favour of EU membership (Doyle and Fidrmuc 2006). We thus include demographic variables in our analysis, measuring age and education. The education data are taken from the Population Project at IIASA (Lutz et al. 2007), which uses definitions and categories consistent over countries and time to facilitate cross-national and time-series comparisons. Precisely, we employ a measure of male secondary education, defined as ‘the proportion of men aged 20-24 years with secondary or higher education of all men aged 20-24’ (Hegre et al. 2013, 255).<sup>13</sup> The age data (*Youth*), in turn, come from the UN World Population Prospects series, which provides age-specific population numbers ‘measured as the percentage of the population aged 15-24 years of all adults aged 15 years and above.’

As to the capabilities argument, that is a state’s administrative capacity to adopt and implement EU rules (or its ability to implement adjustment efforts at the domestic level), factors such as economic development and regime type are among the explanatory variables commonly incorporated in probabilistic models of effective EU conditionality (see Böhmelt and Freyburg 2013; 2015). However, as discussed above, these factors do not meet this study’s requirements. Instead, a common measurement of state capacity in quantitative research is to look at the outcomes of public goods

---

<sup>13</sup> In following Hegre et al. (2013), we also lack data on female secondary education. However, due to our focus on relatively developed countries in Europe (unlike more developing countries outside Europe), there should be a high correlation between male and female secondary education (Breen et al. 2010).

and service delivery, such as the percentage of children enrolled in primary schools, infant mortality rates, or literacy rates. These measures are attractive due to their broad coverage and cross-national comparability, although not without criticism (see Hanson 2015), because they can be attributed to factors other than state capacity, including levels of economic development and the nature of the political regime. Since we do not test the costs and capabilities arguments against each other, but use them to identify a prediction model for candidates' future ability to comply with EU law, we do not perceive this a problem for our study. Considering this discussion, among the traditional measures of state capacity in terms of administrative performance (Bäck and Hadenius 2008), infant mortality rates present a particularly useful composite indicator of the provision of public services (Ross 2006). Taking the data from the UN (2007), infant mortality is defined as the probability of dying between birth and exact age one year, expressed as the number of infant deaths per 1,000 live births.

We also include a time trend, which corrects for temporal dependencies (see Böhmelt and Freyburg 2013). Moreover, the EU has added more policy areas over the years. If the costs to comply increase, because of the addition of further policy areas, compliance might decrease. We log-transform the time-related variable *Year* due to a declining marginal effect on compliance with increasing time (Böhmelt and Freyburg 2013).<sup>14</sup> The also model incorporates fixed effects for a specific enlargement round (i.e., 2004, 2007, and 2013) and future enlargement to account for the spatial dependencies identified in Böhmelt and Freyburg (2015) and the possibility that the specific requirements have become more difficult with each round. These dummies

---

<sup>14</sup> Auffhammer and Carson (2008, 237) recommend against using year fixed effects as 'forecasting model selection criteria punish [this] quite heavily.' Instead, they suggest using a time trend variable, which is our approach.

shall further capture any time-invariant group-specific characteristics and unobserved features of each accession round's requirements, as determined by EU policy (Vachudova 2014).

Finally, we include country-fixed effects that are based on the same rationale at the state level, i.e., capturing unobserved time-invariant unit-level effects that may influence compliance with EU law. In light of these data and methods, particularly the inclusion of country fixed effects, we explicitly build on a 'reduced-form approach,' which assumes that due to the rigid institutional architecture of EU enlargement policy, future policy will continue to develop at roughly the historical pace (see also Schmalensee et al. 1998: 16).<sup>15</sup>

Needless to say, predictions involve – by definition – a degree of uncertainty (see Hegre et al. 2013, 250-251). In the following, however, we demonstrate that the predictors we include can accurately and precisely predict actually observed compliance values, i.e., that our model has little prediction error. Note here that the direction of influence or the statistical significance of the covariates does not matter for this purpose (Ward et al. 2010): it is the precision and accuracy of our final model in making predictions that counts.

## **Empirical Results**

### *In-Sample Prediction*

How accurate are the 'conditional statements about a phenomenon for which the researcher actually has data, i.e., the outcome variable has been observed' (Bechtel

---

<sup>15</sup> In the context of environmental policy-making, Schmalensee et al. (1998: 16) describes this as 'this reduced-form approach of estimation and projection of historical trends amounts to forecasting by „sighting along the data.“ Our estimates thus reflect any relevant historical tightening of environmental standards [or any policies implemented by a third actor, such as the EU in our case], for example, and our projections reflect the (implicit) assumption that such standards would continue to be tightened at roughly the historical pace.'

and Leuffen 2010, 311)? To assess this, we first estimate the baseline model in 1998-2008 with OLS (see Model 1 in Table 2), then calculate the predicted values of this model for that time period, and finally compare the predicted yearly median levels of candidates' compliance with EU accession law using the estimated parameters from the baseline model with the truly observed median compliance between 1998 and 2008. The results are depicted in Figure 2.

While the dashed grey line captures our predicted values as derived from the parameters of our model, the solid black line pertains to the observed values of compliance with the EU *acquis*. The figure shows, on one hand, that the model slightly over-predicts compliance values until about 2000. On the other hand, predicted and observed values are *almost identical as of 2003*, except for the tail of the curve where our model marginally under-predicts compliance. Also note the decline in compliance from 2003 to 2004 – this is driven by those states joining the EU in 2004 dropping out of the sample (as the Commission ceases to provide progress reports with accession); the ‘remaining’ countries in the sample have been, as they were not ready for joining the EU back then, characterized by significantly lower compliance scores. All in all, this figure demonstrates that the predicted values fit the time points of the observed data extremely well.

[Figure 2 here]

To assess the accuracy of this prediction more thoroughly, we use two goodness-of-fit measures: the mean squared prediction error (MSPE) and Theil's U (Theil 1966), which (unlike the MSPE) does not depend on the scale of the data (see also Bechtel and Leuffen 2010). Theil's U is the square root of the ratio between the sum

of squared prediction errors of the baseline model (i.e., Model 1 in Table 2) and the sum of squared prediction errors of a naïve model, i.e., a ‘no-change prediction’ where the level of candidate compliance with EU law in  $t-1$  fully corresponds to the level of compliance in  $t$ . If Theil’s U is larger than 1, the model performs worse than the naïve model; values for Theil’s U smaller than 1 indicate that the ‘theoretically informed model’ performs better than the naïve specification. Moreover, the closer the MPSE is to 0, the more accurate is the model in making predictions. For our model, the MPSE is 0.0065 while Theil’s U stands at 0.723.

Ultimately, therefore, the specification used in the model performs well in predicting candidate states’ compliance with EU law *in-sample*. It remains to be seen, though, how accurately this model predicts candidate states’ compliance when moving to the ‘harder’ test of an out-of-sample prediction confronting the model with ‘new’ data. Put differently, what is the model’s predictive power when trying to correctly predict compliance that is not ‘within the very same set of data that was used to generate the models in the first place’ (Ward et al. 2010, 8)?

### *Out-of-Sample Prediction*

For the out-of-sample prediction, we first use a four-fold cross-validation quasi-experimental setup that was repeated ten times for the baseline model in 1998-2008. We refer the interested reader to Ward et al. (2010, 370) who describe this approach in more detail, but to summarize the main points: cross-validation randomly divides our sample we employed for the baseline model into four segments. We use three segments to estimate the parameters, while the fourth, ‘test-set’ segment (Ward et al. 2010, 370) is retained for assessing the predictive power of the baseline model on the pooled subsets. Therefore, there are three segments of the data to build the model and

create predictions. The remaining (randomly chosen) part of the data is not considered for estimating the model in the first place, and thus treated as if ‘unknown,’ and its mere purpose is for comparing the predicted with the observed values. Again, we calculated the MPSE and Theil’s U for the predictive power, for which we then present the average values over the ten repetitions. The average MSPE for the cross-validation exercise stands at 0.012 while Theil’s U has an average value (across the ten repetitions of the exercise) of 0.966 now. Not surprisingly, the prediction power of the model decreases when confronted with ‘new’ data, although it remains at reasonably high levels.

[Figure 3 here]

Similar to Figure 2, we graphically depict predicted and actual values of candidate states’ compliance with EU law. The difference between Figure 2 and Figure 3 is that the latter extends the period of study to 2016, as we compare the observed values of candidates’ compliance with EU law in 1998-2008 (our ‘observed’ time period) with those ‘unobserved’ values in 2009-2016, i.e., the time period that we have not used for building the model and that we treat as ‘unobserved,’ although we know the true values. The corresponding MSPE is 0.010, while Theil’s U remains below 1 (0.918).

Two conclusions can be derived from this section. First, uncertainty remains and the predictions for the ‘unobserved’ data partition are less accurate than in the case of the in-sample prediction. Two indicators demonstrate this: (1) our goodness-of-fit measures, which both increase and, therefore, show that prediction power decreases, and (2) the comparison between observed (1998-2008) and predicted values (2009-2016) in Figure 3. Having said that, secondly, prediction accuracy is strong, even when confronting the model with new data. In other words, our model improves on



what previous work may have suggested and the naïve model Theil's U refers to as the baseline. Hence, we move on to the core contribution of this article: the out-of-sample forecast of candidates' compliance with EU law in 2017-50 for individual countries and aggregated median predictions.

#### *Out-of-Sample Forecast: 2017-50*

We begin by summarizing the underlying model we use for the forecast. In light of the previous sections, this model is fully based on Model 1 in Table 2 with one exception: we no longer restrict the time period used for building the model's parameters to 1998-2008, but use the entire time period our dependent variable has data for, i.e., 1998-2016. The estimates of Model 2 are virtually identical to those in Model 1. However, the relevant question is how this model predicts the five current and potential candidate states' compliance levels with the EU *acquis* for the future?

[Table 2 here]

To this end, we calculated the predicted values for *EU Accession Compliance* in 1998-2050. Note that these calculations are partly based on data that helped building the model, i.e., we use data that cover 1998-2016. However, all data points after 2016, i.e., 2017-50, are not part of our model as we do not have observed values for candidates' compliance here. This, hence, constitutes the true forecast. We plot these predicted values next to the observed values in the upper-left panel of Figure 4. The solid line signifies the observed values (in 1998-2016), while the two vertical solid lines mark the points in time at which we set thresholds for the in-sample and out-of-sample predictions before (i.e., 2008 and 2016; see above). Further note that we capture the uncertainty inherent in our forecast by including upper and lower bound

confidence intervals for the predictions (upper and lower dashed lines in the upper-left panel of Figure 4).

[Figure 4 here]

We also calculated the predicted values for each of the five current or potential candidate states individually, see Figure 4.<sup>16</sup> The horizontal solid line in these country panels pertains to a reasonable benchmark we set for ‘sufficiently high’ compliance. Specifically, Schimmelfennig and Sedelmeier (2004, 666) acknowledge that the EU might undermine the credibility of its own conditionality if it admits candidates at different levels of preparedness. Testimonies of accession negotiators and coordinators point out that enlargement decisions have been considered for a group of countries as a whole, e.g., the Baltic States or the Visegrad countries rather than for each state individually (see Vassiliou 2007). As a result, particular groups may join the EU when there is a sufficiently large number of candidates who show good compliance (Christoffersen 2007, 32f). We thus assume that all states entering the EU in 2004 (1) had achieved more or less the same level of ‘preparedness’ so that they could jointly assume membership and (2) that this joint level of compliance was also sufficiently high. We, therefore, use the average level of compliance with EU law of all states that joined the EU in 2004 as a benchmark value for future accessions – this benchmark lies at 0.733, according to our data.<sup>17</sup> Again, though, recall that this forecasts states’ ability to comply with the *acquis*, not their willingness and we do not

---

<sup>16</sup> Confidence intervals omitted for the presentation of the point estimates.

<sup>17</sup> In addition, the number of countries that joined the EU in 2004 is significantly larger than those that joined in later accession rounds. Thus, focusing on the 2004 accession round as a benchmark is based on more data points, which lowers measurement error to some degree.

predict actual accession dates (but years in which candidate states might, in principle, be able to comply with EU law).

Several important conclusions can be derived from Figure 4. First, the high prediction power we identified in our model in the previous sections remains to be given. This is illustrated by comparing the predicted and ‘real’ values in 1998-2016 (upper-left panel). Second, overall ability to comply with EU law is supposed to increase over time, regardless of which scenario we look at. Third, and most interestingly, compliance ability varies substantially by country. Despite this variation, though, *only one* of the current and potential EU candidates seems to be able to sufficiently comply with the accession criteria until 2023, while most them are unlikely to be ready for accession before the mid-2030s when taking the 2004 average compliance level as a benchmark. By 2023, only FYR Macedonia is likely to meet the 2004-accession benchmark. Serbia passes the threshold in 2035, Turkey – even after accounting for the recent events pertaining to a democratic backlash, as already reported in the 2016 Commission report – is likely to have the ability to comply with the *acquis* in 2036, while Bosnia-Herzegovina (2050: 0.715) and Albania (2050: 0.722) might find it difficult to meet the set compliance standard even by 2050.

## **Conclusion**

Candidates for EU membership are required to adjust domestic legislation prior to accession for bringing their laws, regulations, and administrative practices in line with the *acquis communautaire*. Knowing in advance which of the (potential) current candidate states are less able to abide by EU regulations over the course of accession is not only of academic interest, but also essential to the EU’s monitoring and enforcement schemes as well as an informed public debate about future EU

enlargement. Yet, thus far, we have known relatively little about the actual accession prospects of current candidate countries, in particular how their compliance ability may develop over the years to come. Previous empirical testing is primarily of an *ex-post* nature and, hence, merely accounts for compliance patterns in the past. We sought to take research on candidates' compliance with the EU accession rules one step further by moving from *ex-post* analysis to predictions and forecasts about likely future compliance.

We specified a model to predict the ability to comply with the EU *acquis* of potential and current candidate countries based on proxies for adjustment costs and administrative capacities, and for which we have observations back to 1998 and projections up to 2050. We used in-sample and out-of-sample techniques to assess the predictive power of that model, before providing out-of-sample forecasts of candidates' compliance with EU law in 2017-50. Our research finds that the empirical model can accurately predict candidate state compliance ability, and it emphasizes that *only one country* of the current and potential EU candidates seems to be able to sufficiently comply with the accession criteria until 2023: FYR Macedonia. Albania or Bosnia-Herzegovina may even face problems in their ability to comply with EU law sufficiently before 2050. These figures appear to lower expectations of further enlargement any time soon. That said, they may still paint a 'too optimistic' picture and the actual future compliance levels of the individual candidate states could well be even weaker than suggested, as our model seems to slightly under-predict compliance rates toward the end of the observation period in Figure 4.

Note that our predictions are based on some partly restrictive assumptions, e.g., that the forecasts for our exogenous predictors turn out to be correct, that the past relationship between our predictors and the probability of EU compliance will

continue to hold in the future, and note that we cannot account for random events or states' unwillingness to comply in the future, even if they could, such as the current AKP government in Turkey (see also Hegre et al. 2013). We conclude, nevertheless, that 'theoretically derived factors that are consistently associated with better predictions should increase the researcher's confidence about their substantive meaningfulness' (Hegre et al. 2017, 116). In fact, assessing the predictive power of empirical models and forecasting state behavior in the future have important implications for theory development and can offer significant benefits for policymakers to foresee candidates' compliance with EU law more accurately. Our study is informative in terms of effectively allocating resources within the EU. That is, the findings strongly suggest that more efforts are necessary if the EU wants better compliance with its rules and regulations for potential member states. What is more, if no political decision will be taken favouring early enlargement despite candidates' non-compliance, our finding that – except for FYR Macedonia – new states are unlikely to be able to join within the next ten years may help to take some heat of the current enlargement debate, which often makes believe that further enlargement is lurking around the next corner.

**Table 1. Covariates Included in the Baseline Model**

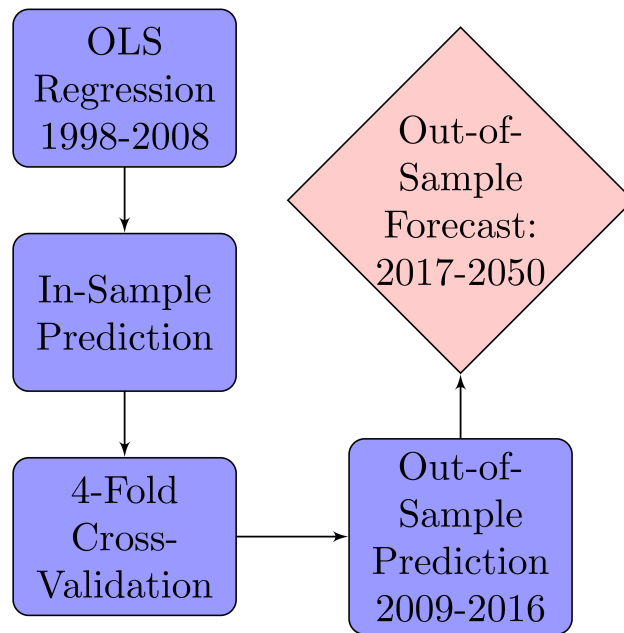
<i>Variable</i>	<i>Description or Source</i>
Year (ln)	Natural logarithm of year
Enlargement 2004	Dummy for respective enlargement rounds, with 1 = member of this round (0 otherwise)
Enlargement 2007	
Enlargement 2013	
Future Enlargement	
Country Fixed Effects	Dummy for respective country, with 1 = country under study (0 otherwise)
Education	International Institute for Applied Systems Analysis
Relative Youth Size Cohort	UN World Population Prospects 2006
Infant Mortality Rate	UN World Population Prospects 2006

**Table 2. Baseline Model of EU Accession Compliance**

	<b>Model 1</b>	<b>Model 2</b>
Year (ln)	0.287 (0.022)	0.296 (0.021)
Enlargement 2004	0.246 (0.138)	0.244 (0.160)
Enlargement 2007	0.122 (0.110)	0.240 (0.140)
Future Enlargement	0.116 (0.069)	0.055 (0.050)
Youth	0.006 (0.020)	-0.025 (0.015)
Education	2.044 (1.214)	0.689 (0.814)
Infant Mortality Rate	0.017 (0.010)	0.024 (0.008)
Constant	-2.298 (1.244)	-0.820 (0.843)
<i>Observations</i>	<i>120</i>	<i>164</i>
<i>Country Fixed Effects</i>	<i>Yes</i>	<i>Yes</i>
<i>Prob &gt; F</i>	<i>0.000</i>	<i>0.000</i>
<i>Adj. R<sup>2</sup></i>	<i>0.853</i>	<i>0.813</i>

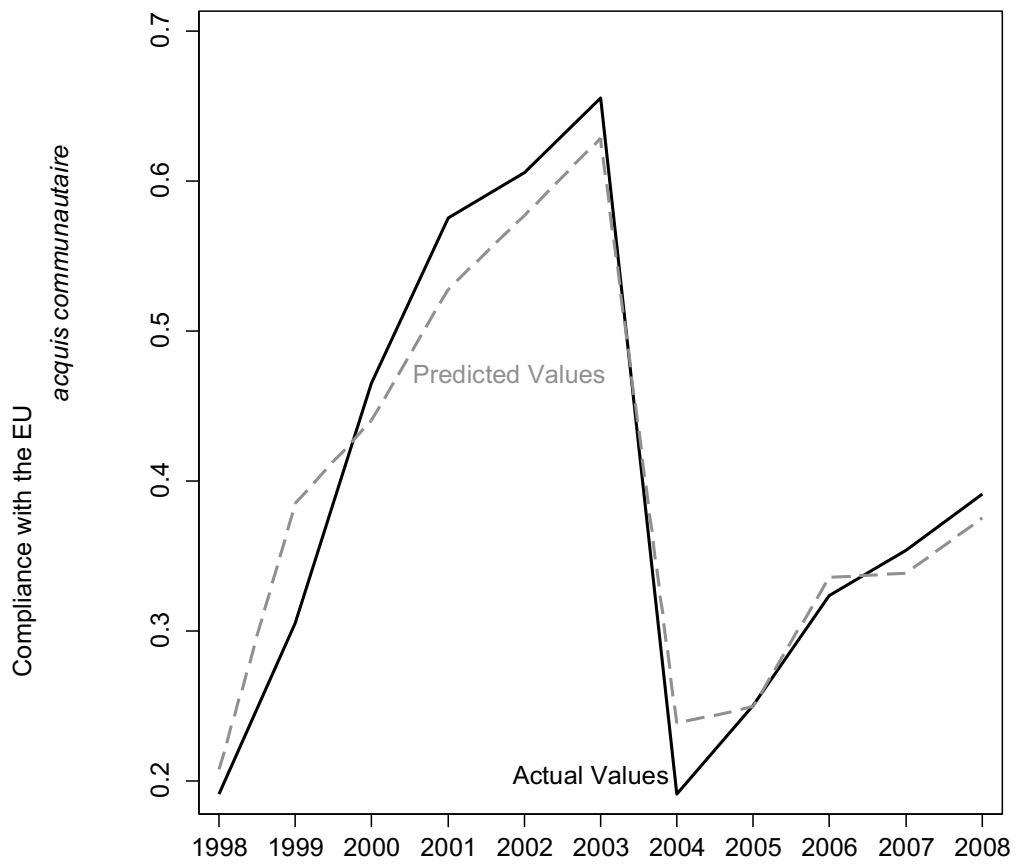
*Note:* Standard errors in parentheses. Because of collinearity, the 2013 enlargement round and the 2004 enlargement round are omitted in in Model 1 and Model 2, respectively.

**Figure 1. Multiple-Step Forecasting Procedure**





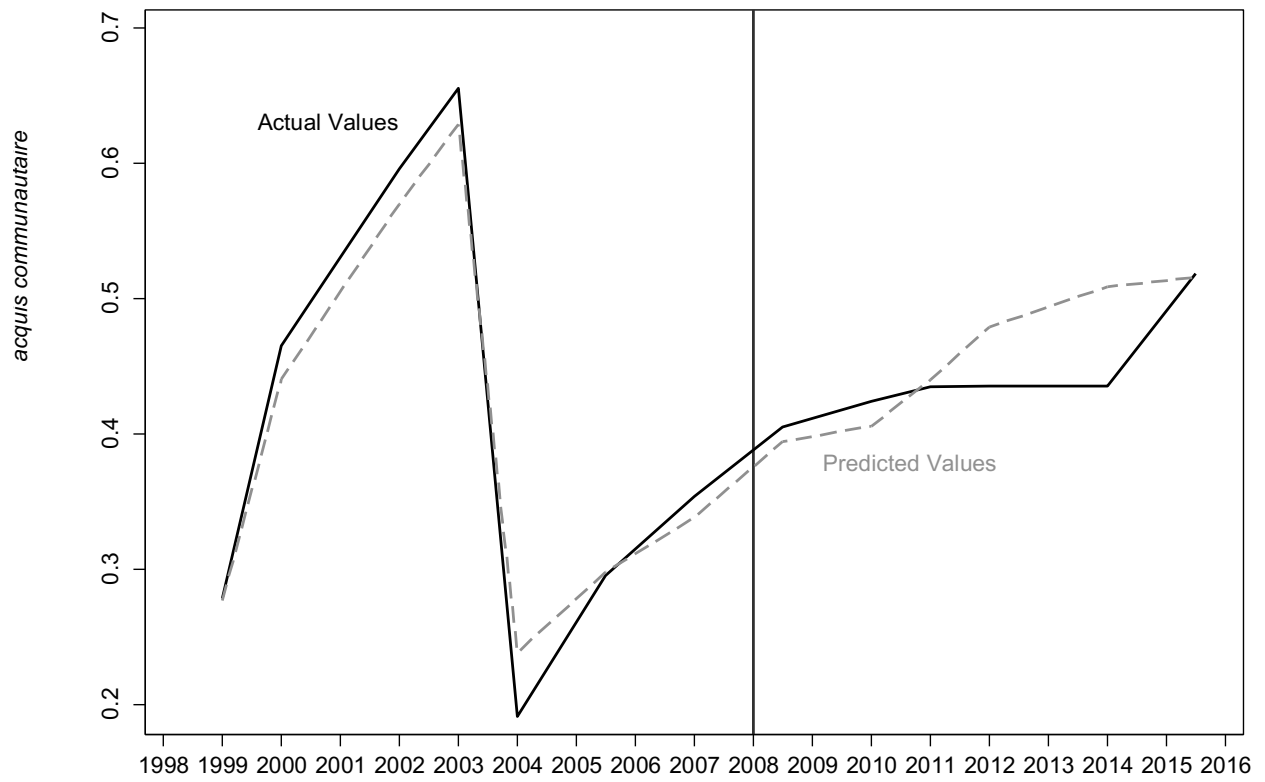
**Figure 2. Median Levels of Compliance with EU Accession Rules – In-sample Prediction**



*Note:* Predicted values are indicated by the dashed line and actual values by the solid line.

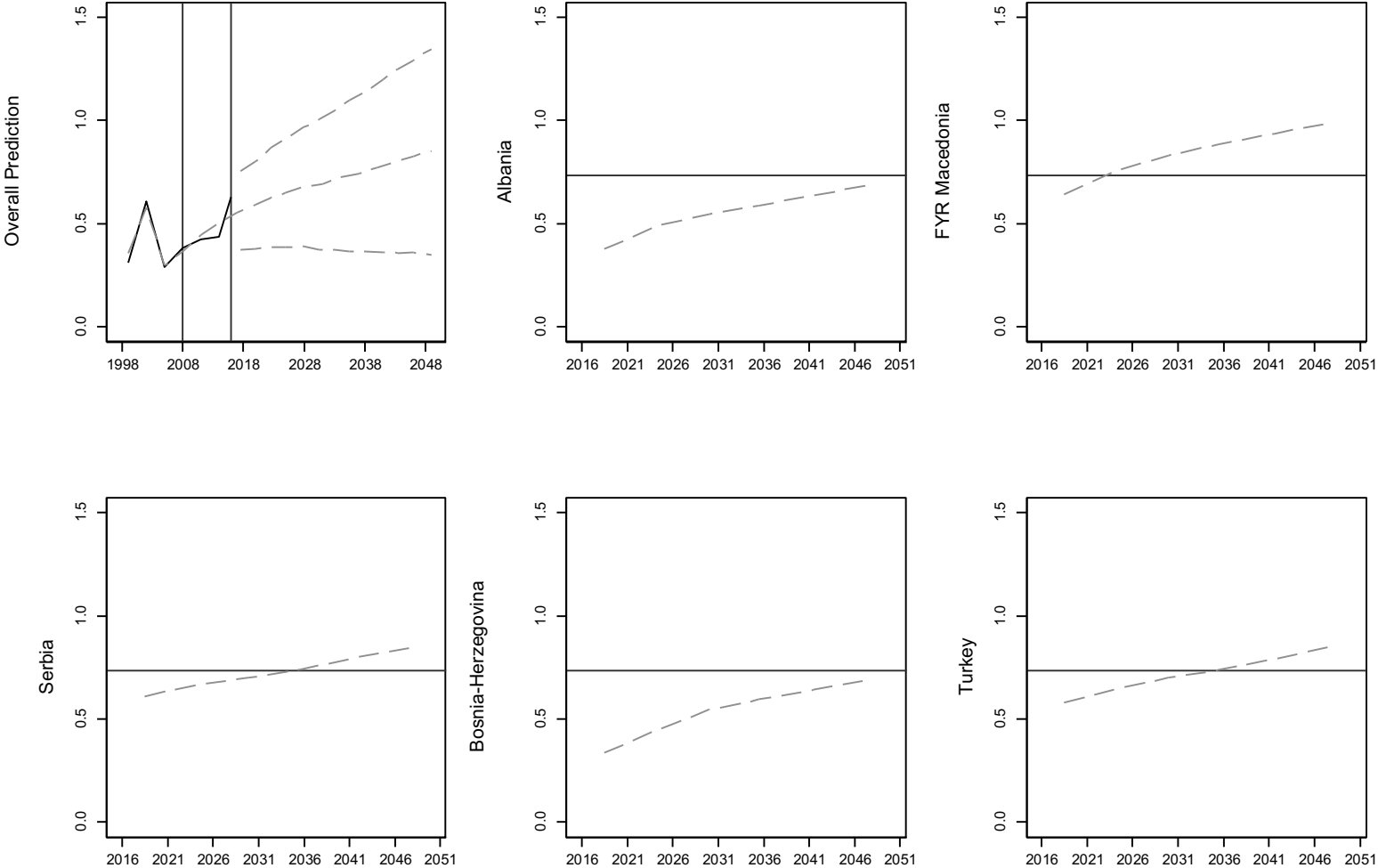
**Figure 3. Median Levels of Compliance with EU Accession Rules – Out-of-Sample**

**Prediction**



*Note:* Predicted values are indicated by the dashed line and actual values by the solid line.

**Figure 4. Median Levels of Compliance with EU Accession Rules**



*Note:* Predicted values are indicated by the dashed line and actual values by the solid line. Solid horizontal line marks benchmark compliance value (2004 accession round). Vertical solid lines in upper-left panel pertain to the points in time at which we set thresholds for the in-sample and out-of-sample predictions (i.e., 2008 and 2016). Upper and lower bounds of 90 percent confidence interval are included in upper-left panel in 2017-50 as well.

## References

- Auffhammer, M. and Carson, R. (2008) 'Forecasting the Path of China's CO<sub>2</sub> Emissions Using Province-Level Information', *Journal of Environmental Economics and Management* 55(3): 229-47.
- Bäck, H. and Hadenius, A. (2008) 'Democracy and State Capacity: Exploring a J-Shaped Relationship', *Governance* 1(1): 1-24.
- Bechtel, M. and Leuffen, D. (2010) 'Forecasting European Union Politics: Real-Time Forecasts in Political Time Series Analysis', *European Union Politics* 11(2): 309-27.
- Beck, N., King, G. and Zeng, L. (2000) 'Improving Quantitative Studies of International Conflict: A Conjecture', *American Political Science Review* 94(1): 21-35.
- Böhmelt, T. and Freyburg, T. (2013) 'The Temporal Dimension of the Credibility of EU Conditionality and Candidate States' Compliance with the *acquis communautaire*, 1998-2009', *European Union Politics* 14(2): 250-72.
- Böhmelt, T. and Freyburg, T. (2015) 'Diffusion of Compliance in the 'Race towards Brussels?' A Spatial Approach to EU Accession Conditionality', *West European Politics* 38(3): 601-26.
- Börzel, T. et al. (2010) 'Obstinate and inefficient: Why member states do not comply with European law.' *Comparative Political Studies* 43(11): 1363-90.
- Breen, R., Luijkx, R., Müller W. and Pollak, R. (2010) 'Long-Term Trends in Educational Inequality in Europe: Class Inequalities and Gender Differences', *European Sociological Review* 26(1): 31-48.

- Christoffersen, P. (2007) 'The Preparation of the Fifth Enlargement', in G. Vassiliou (ed.) *The Accession Story. The EU from Fifteen to Twenty-Five Countries*, Oxford: Oxford University Press, pp. 24-33.
- Desmarais, B. and Cranmer, S. (2013) 'Forecasting the Locational Dynamics of Transnational Terrorism: A Network Analytic Approach', *Security Informatics* 2(8): 1-13.
- Doyle, O. and Fidrmuc, J. (2006) 'Who Favors Enlargement? Determinants of Support for EU Membership in the Candidate Countries' Referenda', *European Journal of Political Economy* 22(2): 520-43.
- European Commission (2016) '2016 Communication on EU Enlargement Policy', *COM(2016)715 Final*, 9 November, Brussels: Commission of the European Communities, available at [https://ec.europa.eu/neighbourhood-enlargement/sites/near/files/pdf/key\\_documents/2016/20161109\\_strategy\\_paper\\_en.pdf](https://ec.europa.eu/neighbourhood-enlargement/sites/near/files/pdf/key_documents/2016/20161109_strategy_paper_en.pdf) (accessed April 10, 2017).
- Gabel, M. (1998) 'Public Support for European Integration: An Empirical Test of Five Theories', *Journal of Politics* 60(2): 333-64.
- Giandomenico, J. (2009) 'Path Dependency in EU Enlargement: Macedonia's Candidate Status from a Historical Institutionalist Perspective', *European Foreign Affairs Review* 14: 89-112.
- Goertz, G. (2006) *Social science concepts: A user's guide*. Princeton University Press, Princeton, NJ.
- Goldstone, J. et al. (2010) 'A Global Model for Forecasting Political Instability', *American Journal of Political Science* 54(1): 190-208.
- Hanson, J. (2015) 'Democracy and State Capacity: Complements or Substitutes?', *Studies in Comparative International Development* 50(3): 304-30.

- Hegre, H., Karlsen, J., Nygård, H., Strand, H., Urdal, H. and Karlsen, J. (2013) 'Predicting Armed Conflict, 2010-2050', *International Studies Quarterly* 57(2): 250-70.
- Hegre, H., Metternich, N., Nygård, H. and Wucherpfennig J. (2017) 'Introduction: Forecasting in peace research', *Journal of Peace Research* 54(2): 113-24.
- Hille, P. and Knill, C. (2006) "'It's the Bureaucracy, Stupid.": The Implementation of the Acquis Communautaire in EU Candidate Countries, 1999-2003', *European Union Politics* 7(4): 531-52.
- Kelley, J. (2006) 'New Wine in Old Wineskins: Promoting Political Reforms through the New European Neighbourhood Policy', *Journal of Common Market Studies* 44(1): 29-55.
- Lutz, W., Goujon, A., Samir, H.C. and Sanderson, W. (2007) 'Reconstruction of Population by Age, Sex, and Level of Education Attainment of 120 Countries for 1970-2000', *Vienna Yearbook of Population Research* 5(1): 193-235.
- Mattli, W. and Plümper, T. (2002) 'The Demand-Side Politics of EU Enlargement: Democracy and the Application for EU Membership', *Journal of European Public Policy* 9(4): 550-74.
- O'Brennan, J. (2014) "'On the Slow Train to Nowhere?": The European Union, 'Enlargement Fatigue' and the Western Balkans', *European Foreign Affairs Review* 19(2): 221-42.
- O'Neill, B., Balk, D., Brickman, M., and Ezra, M. (2001) 'A Guide to Global Population Projections', *Demographic Research* 4(8): 203-88.
- Ross, M. (2006) 'Is Democracy Good for the Poor?', *American Journal of Political Science* 50(4): 860-74.

- Schimmelfennig, F. (2008) 'EU Political Accession Conditionality After the 2004 Enlargement: Consistency and Effectiveness', *Journal of European Public Policy* 15(6): 918-37.
- Schimmelfennig, F. (2009) 'Entrapped Again: The Way to EU Membership Negotiations with Turkey', *International Politics* 46(4): 413-31.
- Schimmelfennig, F. and Scholtz, H. (2008) 'EU Democracy Promotion in the European Neighbourhood: Political Conditionality, Economic Development, and Transnational Exchange', *European Union Politics* 9(2): 187-215.
- Schimmelfennig, F. and Sedelmeier, U. (2004) 'Governance by Conditionality? EU Rule Transfer to the Candidate Countries of Central and Eastern Europe', *Journal of European Public Policy* 11(4): 661-79.
- Schmalensee, R., Stoker T. and Judson R. (1998) 'World Carbon Dioxide Emissions: 1950-2050', *Review of Economics and Statistics* 80(1): 15-27.
- Schneider, G., Gleditsch, N.P and Carey, S. (2010). 'Exploring the past, anticipating the future', *International Studies Review* 12(1): 1-7.
- Schneider, G., Gleditsch, N.P. and Carey, S. (2011) 'Forecasting in International Relations: One Quest, Three Approaches', *Conflict Management and Peace Science* 28(1): 5-14.
- Steunenberg, B. and Dimitrova, A. (2007) 'Compliance in the EU Enlargement Process: The Limits of Conditionality', *European Integration Online Papers* 11(5): 1-18.
- Szymański, A. (2013) '*EU Enlargement Policy – Critical Review of the Theoretical Research Framework*', Typescript: University of Warsaw.
- Theil, H. (1966) *Applied Economic Forecasts*, Amsterdam: North-Holland Publishing.

- Thelen, K. and Steinmo, S. (1992) *Structuring Politics. Historical Institutionalism in Comparative Analysis*, Cambridge: Cambridge University Press.
- Thomson, R. and Hosli, M. (2006) 'Who Has Power in the EU? The Commission, Council and Parliament in Legislative Decision-Making', *Journal of Common Market Studies* 44(2): 391-417.
- Toshkov, D. (2008) 'Embracing European Law Compliance with EU Directives in Central and Eastern Europe', *European Union Politics* 9(3): 379-402.
- Toshkov, D. (2010) *Taking Stock: A Review of Quantitative Studies of Transposition and Implementation of EU Law*, Typescript: University of Leiden.
- Tucker, J., Pacek, A. and Berinsky, A. (2002) 'Transitional Winners and Losers: Attitudes toward EU Membership in Post-Communist Countries', *American Journal of Political Science* 46(3): 557-71.
- Tverdova, Y. and Anderson, C. (2004) 'Choosing the West? Referendum Choices on EU membership in East-Central Europe', *Electoral Studies* 23(2): 185-208.
- United Nations (2007) *World Population Prospects. The 2006 Revision*, New York: UN.
- Vachudova, M. (2007) 'Historical Institutionalism and the EU's Eastern Enlargement', in K. McNamara and S. Meunier (eds) *Making History: European Integration and Institutional Change at Fifty*, Oxford: Oxford University Press, pp. 105-22.
- Vachudova, M. (2014) 'EU Leverage and National Interests in the Balkans: The Puzzles of Enlargement Ten Years On', *Journal of Common Market Studies* 52(1): 122-38.
- Vassiliou, G. (2007) *The Accession Story. The EU from Fifteen to Twenty-Five Countries*, Oxford: Oxford University Press.



Ward, M., Greenhill, B. and Bakke, K. (2010) 'The Perils of Policy by  $p$ -Value: Predicting Civil Conflicts', *Journal of Peace Research* 47(4): 363-75.

## APPENDIX

### Appendix 1. Covariates Used in Existing Quantitative Research

<i>Variable</i>	<i>Effect (significance)</i>	<i>Reference</i>
<u>Costs</u>		
Political constraints index	positive significant	Hille/Knill 2006
(Henisz's Polcon III index)	~ zero	Böhmelt/Freyburg 2013
	~ zero	Böhmelt/Freyburg 2015
Government's position	positive	Hille/Knill 2006
(Hille/Knill based on Benoit and Laver 2006)	negative significant	Böhmelt/Freyburg 2013
	~ zero	Böhmelt/Freyburg 2015
Political liberalization	~ zero	Hille/Knill 2006
(Polity IV)	positive significant	Böhmelt/Freyburg 2013
	~ zero	Böhmelt/Freyburg 2015
Economic liberalization	positive significant	Böhmelt/Freyburg 2013
(Heritage Foundation Index)	~ zero	Böhmelt/Freyburg 2015
<u>Capabilities</u>		
Government expenditures, p.c.	~ zero	Hille/Knill 2006
(CIA factbook; WB)	~ zero	Böhmelt/Freyburg 2013
	~ zero	Böhmelt/Freyburg 2015
Gross domestic product, p.c.	positive	Hille/Knill 2006
(Commission reports; WB)	negative significant	Böhmelt/Freyburg 2013
	~ zero	Böhmelt/Freyburg 2015
Bureaucratic strength	positive significant	Hille/Knill 2006
(WB governance Index)	~ zero	Böhmelt/Freyburg 2013
	~ zero	Böhmelt/Freyburg 2015

---

Other

Pro-enlargement presidency	~ zero	Böhmelt/Freyburg 2013
Membership probability	positive significant	Böhmelt/Freyburg 2013
IGO membership	positive significant	Böhmelt/Freyburg 2015

(Pevehouse et al. 2004)

---

*Note:* ~ zero if  $p < 0.05$  and  $\beta \leq 0.1$ ; data source in brackets. Reported are Model 2 in Hille and Knill (2006), Model 3 in Böhmelt and Freyburg (2013), and Model 4 in Böhmelt and Freyburg (2015).

## Appendix 2. Countries and Years Considered for Analysis

<i>Country</i>	<i>Year Coverage</i>	<i>Application</i>	<i>Candidate Status</i>	<i>Negotiations Start</i>	<i>Negotiations End</i>	<i>Accession Treaty</i>
Albania	2003-2016 [2050]	2009	2014	pending	pending	pending
Bosnia-Herzegovina	2004-2016 [2050]	2016	pending	pending	pending	pending
Bulgaria	1998-2006	1995	1997	2000	2004	2005
Croatia	2002-2012	2003	2004	2005	2011	2011
Cyprus	1998-2003	1990	1997	1998	2002	2003
Czech Republic	1998-2003	1996	1997	1998	2002	2003
Estonia	1998-2003	1995	1997	1998	2002	2003
Hungary	1998-2003	1994	1997	1998	2002	2003
Latvia	1998-2003	1995	1997	2000	2002	2003
Lithuania	1998-2003	1995	1997	2000	2002	2003
Macedonia	2002-2016 [2050]	2004	2005	pending	pending	pending
Malta	1999-2003	1990	1997	2000	2002	2003
Poland	1998-2003	1994	1997	1998	2002	2003
Romania	1998-2006	1995	1997	2000	2004	2005
Serbia	2006-2016 [2050]	2009	2012	2014	pending	pending
Slovakia	1998-2003	1995	1997	2000	2002	2003
Slovenia	1998-2003	1996	1997	1998	2002	2003
Turkey	1998-2016 [2050]	1987	1999	2005	pending	pending

*Note:* Kosovo and Montenegro are excluded due to lack of data for covariates. Also, Kosovo's independence has not yet been recognized by all EU member states and Montenegro became independent only in 2006. We also omit Iceland as its government effectively withdrew its application in March 2015; the country had also incorporated about two-thirds of the *acquis* chapters into its legislation already prior to the start of the accession procedure in 2009 through membership in the European Economic Area (EEA), the Schengen Area, and the European Free Trade Association (EFTA).