Supplementary material

A review of the empirical strategies for the study of environmental agreements

Francesco S. Bellelli, Ashar Aftab, Riccardo Scarpa*
21st June 2021

The act of ratification refers to a specific agreement, originates by a distinct country and occurs at a fixed point in time. In essence, it is qualified by three dimensions: the ratifying country, the ratified treaty, and the year of ratification. According to their methodology, the empirical research emphasised different combinations of these dimensions, looking from different angles at the same phenomenon. Over time the methodological approaches followed a process of refinement, gradually attempting to include all three dimensions and leading to more general conclusions on ratification. We broadly distinguish between three empirical approaches to empirical ratification analysis: *i*) ratification counts, *ii*) survival analysis for single agreements, and *iii*) pooled survival analysis. In this appendix we describe each of these approaches, their applications, characteristics and limitations. Moreover, as a reference to the reader, we provide at the end of each section a table summarising the sample and models of surveyed studies adopting that approach (Tables 1, 2 and 4).

1 Ratification counts

The first step in the analysis of ratification is to find an appropriate way of "measuring" the ratification behaviour of countries. This is usually done by looking at the ratification status of one or more treaties at a specific moment in time and counting the number of treaties ratified by every country. If only a single agreement is involved, the maximum count is 1 and the variable is binary. If more than one agreement is studied, the variable represents the total number of treaties in which the country decided to participate. We call this type of variable a *ratification count*, to distinguish it from the survival data employed in later studies (e.g. Fredriksson et al. 2007, Bernauer et al. 2010). Count studies focus primarily on the difference in the number of ratifications among countries, while the evolution in time is generally ignored. Almost all of the studies measuring ratification in a "count" fashion adopt a cross-sectional approach. Egger et al. (2011 and 2013) and Davies and Naughton (2014) are the only panel studies using count data (see Table 1).

*Centre for Environmental and Energy Economics, Durham University Business School, University of Durham, Durham DH1 3LB, United Kingdom. Dr. Bellelli: (e-mail: f.s.bellelli@dunelm.org.uk); Dr. Aftab (e-mail: ashar.aftab@durham.ac.uk); Prof. Scarpa (e-mail: riccardo.scarpa@durham.ac.uk).

Depending on whether the study covers a single treaty or multiple treaties, the researcher deals with two different types of data. Hence, different empirical strategies are used. When the study is limited to a single agreement the ratification data is represented by a binary variable: this type of analysis is approached with binary regression. When multiple agreements are studied, the dependent variable is the total number of ratifications: this data is approached either with a linear regression or with regression techniques for count data.

1.1 Single agreements: the binary outcome

As mentioned earlier, the simplest case of ratification counts is when only a single agreement is studied (e.g. Beron et al. 2003, Murdoch et al. 2003) or when agreements are modelled individually (e.g. Congleton 1992, Neumayer 2002a). In these cases the dependent variable is binary because the maximum count is 1.

Binary ratification choices are tackled with a binomial regression to study how differences among countries affect the odds of ratification. This approach has been implemented in numerous studies. Congleton (1992), Almer and Winkler (2010) and Neumayer (2002b) use it to model the signature of environmental agreements. Murdoch et al. (2003) and Beron et al. (2003) study the ratification of two different protocols by 25 and 89 countries, respectively. Additional work is conducted by Almer and Winkler (2010) and Fredriksson and Ujhelyi (2006), both investigating the ratification of the Kyoto Protocol by circa 170 countries.

These papers study exclusively one agreement, raising the question of how the results can be generalised beyond the single case. They fulfil a descriptive purpose and offer little insight into the general process of ratification. This is probably the main weakness of this approach. Frank (1999) and Neumayer (2002a and 2002b) attempt to expand the range of this type of studies by modelling several agreements in parallel. However, the results are still based on individual models for each treaty and the dimension of the sample is relatively small. In principle, the approach could be extended to several agreements by using ratification dummies for treaty-country dyads; however, this strategy has never been implemented.

1.2 Multiple agreements: counting ratifications

Whenever the number of ratifications are counted for two or more agreements, we are effectively dealing with count data. This type of data has been fitted with count models and — more commonly, but less appropriate given the positive and discrete nature of the count variable — with linear models. Ratification counts are an easy way to expand the base of treaties included in the analysis because with this approach less information is required compared to approaches based on survival analyses. Recchia (2002) covers 15 environmental treaties, Roberts et al. (2004) 22 and Seelarbokus (2014) reaches 110 agreements. This is considerably more than any other cross-sectional study using binary variables.

Nonetheless, the simple count of ratifications gives rise to a likely misleading variable if the objective is to define a country's appetite for international cooperation. The implicit assumption is that more ratified treaties lead to stronger environmental commitments. This assumption is debatable because environmental agreements are profoundly different among them: simply adding treaties up without adequate weighing is like summing 'grapes and melons'. The number of agreements that are ratified is unlikely to be proportional to either the environmental commitment of the country or representative of its engagement in the international arena of environmental cooperation. To a large extent, the number of ratified treaties is just a reflection of the number of treaties the country can access. To this end, to evaluate a country's opportunity set, it is critical to know the number of neighbours and environmental issues in which a country could be involved. As an illustration, Kiribati is an insular state in the pacific, despite its interest in preserving the environment, it undoubtedly ratifies fewer agreements than Indonesia, a big state with several neighbours and a rich natural asset. This is due to the massively different opportunity set between the two countries, more than to the country's appetite for international cooperation. Regrettably, neither Seelarbokus (2014) nor Roberts et al. (2004) controlled for these important factors.

We suggest that ratification rate would be a better measure than the mere count of the number of ratifications. Of course, this entails identifying the potential ratifiers of each treaty, a practice implemented for the first time by Bellelli et al. (2020). There has also been some attempt to use score systems instead of simple ratification counts. They usually work by assigning points for signatures and ratifications (Recchia, 2002) or by weighting the number of ratifications by the total number of ratifiers (Roberts, 1996). It unclear what these indices could teach about the ratification of environmental agreements. In general, score systems tend to obfuscate the results, making the relationship between variables opaque.

A less obvious consequence of ratification counts is that the connection between the ratifying country and the ratified treaty can be maintained only if each agreement is studied individually. That is to say, if we sum all the ratifications of a country, we would not be able to tell which types of treaty it has ratified, except in the trivial cases in which it has ratified none or all of them. This feature is a severe limitation to using count variables because it does not allow researchers to study how the design of the treaty affects ratification. The characteristics of a treaty can be accounted for only by studying a cross-section of treaties and counting the number of ratifications it has received, just as in Bernauer et al. (2013b). However, this implies that it would now be impossible to know what country ratified, and to consequentially investigate the role of a country's characteristics in the process. In essence, with such dependent variables, there is a trade-off between studying the characteristics of the country or the treaty.

1.3 Cross-sectional approach

Virtually all count studies adopt a cross-sectional approach; the only ones employing panel data are Egger et al. (2011 and 2013) and Davies and Naughton (2014).

The first problem encountered by the researcher applying cross-sectional approaches is to choose the right cut-off date. Since data is right-censored by construction, ratifications that took place after the cut-off date are ignored. For more recent treaties this could lead to misleading results because the selection of the observation point can arbitrarily influence the results. The choice of the cut-off date is a common problem in studies using ratification counts. For instance, Beron et al. (2003) allow only three years for the ratification of the Montreal Protocol (1987), while in Murdoch et al. (2003) the observations on the Helsinki Protocol (1985) are taken after five years: in both cases, most of the ratifications were not yet deposited by the time the analysis took place. Congleton (1992), Neumayer (2002a) and Neumayer (2002b) study recent environmental agreements but mitigate the problem by focusing on the act of signature — which typically takes place during the first year of the treaty — instead of the act of ratification. The problem is particularly serious for studies with large samples of treaties because different treaties are exposed to ratification processes for different lengths of time. All the studies mentioned in this section fail to address this issue, except for Bernauer et al. (2013b), who account for the exposition factor by using a negative binomial model.

An additional problem of cross-sectional studies is that they ignore the temporal dynamics of ratification. Many domestic policy and institutional factors are likely to influence the timing rather than the occurrence of ratification. For example, Spilker and Koubi (2016) analyse how different domestic voting requirements for the ratification of international treaties influence the likelihood of ratification. It is reasonable to expect that complex or stricter requirements would make the adoption of a treaty not just more laborious, but also slower. The empirical results support this view, countries that require a supermajority in parliament for the approval of treaties are slower and less likely to ratify environmental agreements. Moreover, if time is ignored, it is also impossible to discern the order in which different countries decide to join a treaty, which could provides useful evidence of the diplomatic interactions at play (Almer and Winkler, 2010).

1.4 Panel approach

The obvious solution to the omission of time effects is to use stacked cross-sections to create a panel dataset, this approach has been attempted by Davies and Naughton (2014) and Egger et al. (2011 and 2013). Davies and Naughton (2014) study participation in 110 environmental agreements by 139 countries over 20 years (1980–1999). The dependent variable is a count of ratifications. The study has a very robust methodological approach; the main weakness of the paper being the use of count data. Davies and Naughton (2014) build a spatial model and experiment with different estimators (notably 2SLS and GMM). They use an instrumental variable approach to address endogeneity in one of the variables (Foreign Direct Investment, or FDI) and include country and year fixed effects to account for fixed unobserved factors.

Unfortunately, the use of count data does not serve well the aim of Davies and Naughton (2014). Their objective is to assess the influence of FDI on environmental policies and determine whether ratification is sensitive to the participation decision of neighbouring countries. The problem with choosing the count of ratifications as dependent variable is that it obfuscates interactions between countries. Does the fact that foreign nations ratified a higher number of agreements mean that they had an impact on the domestic ratification choices? How do we know they ratified the same agreement? Could it not reflect the fact that a larger number of agreements have been agreed and are open to ratification? The research question cannot be properly answered because count data does not allow to compare ratification choices within the same agreements. Indeed, count data lumps together the ratifications of different treaties, thus losing information on which specific treaties was ratified by every given nation.

Similarly to the previous study, the dependent variable in Egger et al. (2011 and 2013) is the number of agreements in which a country participates at any given point in time. The definition of "participation" is not clear in the 2011's paper: it appears that a country is considered to be a participant if it either signs or ratifies an agreement, regardless of which. However, in Egger et al. (2013) reference is made to the act of ratification. Their data covers the ratification status of around 350 treaties for 105 countries — of which only 17 Less Developed Countries (LDCs), suggesting that there could be sampling bias. The same control variables and methodological approach are used in both papers. In both Egger et al. (2011) and Egger et al. (2013) a dynamic feedback model for count data with lagged dependent variable is used to model the number of ratifications. The main difference is that in Egger et al. (2013) a separate model is estimated for different clusters of environmental treaties (atmosphere, land, sea, biodiversity protection and hazardous waste).

The main downside of a panel approach with count data is that it does not allow the analyst to escape the trade-off between country and treaty characteristics. If the dependent variable is the number of ratified treaties by the country at a given point in time, then it is not possible to know what treaty the country has ratified. Consequently, the characteristics of the treaties cannot be used to explain its ratification. In the same way, if the focus is on the number of ratifications received by the treaty at time t, then it is not possible to discern which country ratified and take into account the characteristics of the country to explain the ratification choices.

Table 1: Studies modelling ratification as a count or binary variable

Paper	Sample	Dependent variable	Model
Congleton (1992)	118 countries, Vienna Convention (1985) and Montreal Protocol (1987).	Signature by 1989, binary variable.	Logistic regression.
Roberts (1996)	145 countries, 9 environmental agreements.	Weighted number of ratifications between 1963–1987.	Linear regression.
Frank (1999)	Unspecified number of treaties, between 41 and 114 countries depending on time window.	Total number of treaties ratified by a country over 4 time windows.	4 latent variable regressions.
Neumayer (2002a)	6 agreements, maximum of 175 countries.	*i*) Survival data for ratification of 3 agreements. *ii*) Binary variable for the signature of 3 other agreements by 2000.	*i*) Cox PH models for 3 treaties with high ratification rate. *ii*) Probit models for the signature of 3 recent agreements for which ratification process is at its beginning.
Neumayer (2002b)	4 agreements with non-universal ratification, maximum of 175 countries.	Binary variable for the signature (ratification for the Montreal Protocol, 1987) by 2000.	*i*) Probit regressions for single agreements. *ii*) Ordered probit for joint regression (from 0 to 4).
Recchia (2002)	15 global environmental agreements, 19 democracies.	Country score calculated by assigning 3 points for each ratified agreements and 1 point for signature.	Linear regression.

Table 1: Studies modelling ratification as a count or binary variable (continued)

Paper	Sample	Dependent variable	Model
Beron et al. (2003)	Montreal Protocol (1987), 89 countries.	Binary variable for ratification by 1990.	Probit with spatial lag. Weighting matrix based on bilateral trade.
Murdoch et al. (2003)	Helsinki Protocol (1985), 25 European countries.	Binary variable for ratification by 1990.	Probit model.
Roberts et al. (2004)	22 agreements, 192 countries.	Index based on the number of ratifications between 1947-1999.	Linear regression.
Almer and Winkler (2010)	Kyoto Protocol (1997), 165 countries.	*i*) Binary variable for the signature and *ii*) ordered variable for the ratification of the protocol.	A latent variable approach is used for the binary variable (signature yes/no) and an ordered response model for ratification (ratified in period 1, 2 or 3).
Egger et al. (2011)	353 agreements between 1960 and 2006, 105 countries.	Number of agreements in which a country is participating.	Dynamic panel linear feedback model for count data, estimated with GMM.
Egger et al. (2013)	110 countries, more than 212 agreements signed between 1960 and 2006	Number of participation in agreements by country.	Dynamic panel linear feedback model for count data, estimated with GMM. A model is estimated for every cluster of environmental treaties (atmosphere, land, sea, biodiversity protection, hazardous waste).
Bernauer et al. (2013b)	200 agreements.	Total number of ratifications received by each agreement by 2006.	Negative binomial regression.
Davies and Naughton (2014)	110 environmental agreements, 139 countries over 1980-1999.	Number of agreements ratified.	Panel count spatial model with weights based on bilateral distance. Country and year fixed effects. Estimated with GMM-IV and 2SLS.
Seelarbokus (2014)	110 environmental agreements, 108 countries.	Number of treaties ratified or signed by each country.	Linear regression.

2 Survival analysis

So far, we discussed the studies that "measure" ratification behaviour by counting ratification acts by countries. While this approach is the most common in earlier studies, later studies shifted towards the use of survival analysis. Survival analysis derives its name from the epidemiological background of the technique; it is used to study the probability of occurrence of an event at a specific point in time. Following this approach, the ratification of environmental agreements is characterised by two dimensions. The first is whether or not ratification takes place — the *occurrence*. The second is the *timing* to ratification. Hence, compared to ratification counts, survival data incorporates additional information regarding the variation of timing across countries. In this section, we only review those studies that either focus on single treaties or model treaties individually (see Table 2). This methodology can be extended to a plurality of agreements as described in the next section.

The first application of survival models to the ratification of international agreements was by Fredriksson and Gaston (2000), where the authors argue that country's environmental commitment drives the speed of ratification. This relationship, however, also dependent on frictions encountered during the internal procedures of ratification, which vary across different institutional designs. Unfortunately, Fredriksson and Gaston (2000) failed to account for such aspects in the timing of ratification. In subsequent research, it was realised that time to ratification is a better dependent variable than the simple occurrence of ratification, because many factors result in changes in timing rather than occurrence (von Stein 2008, Fredriksson and Ujhelyi 2006). This notion is particularly important in works focusing on the role of political and economic variables. In fact, at the margin, a slightly more complex bureaucratic system, or a small increment in the pressure of environmental groups, are more likely to affect the timing rather than completely reversing the outcome of ratification.

2.1 The information value of timing of ratification

Compared to ratification counts, survival analysis allows taking advantage of the information carried by the timing of ratification. This added dimension allows researchers to expand the scope of the empirical analysis to address new types of questions.

Focus on timing of ratification allows researchers to gather information by observing the behaviour of countries over a specific observation period. Such period starts when the treaty is opened to the debate leading to ratification. From that moment, the country is considered *at risk* of ratification. Ratification by different nations is then tracked throughout time until the cut-off (censoring) year. Ratifications that take place after the censoring year are ignored. Nevertheless, survival analysis is designed to cope with right-censoring. Estimation results are unbiased as long as the assumption of non-informative censoring is satisfied. That is to say, whenever the ratification process and the observation cut-off date are independent.

The advantage of the survival approach is that it can measure ratification over an additional dimension: that of time. Neumayer (2002b) uses this approach to his advantage as he observes that a cross-sectional ratification count study is unable to detect variability within almost-universally ratified treaties. He applies the technique to the Montreal Protocol (1987), CITES (1973) and the Biodiversity Convention (1992), which, by the time the analysis was conducted, had already been ratified by a very large number of nations. In general, survival analysis is a superior approach for universally ratified treaties because it takes advantage of the heterogeneity in the time dimension, while a cross-sectional count study fails to capture any differences in the ratifications when almost all countries have ratified. Survival analysis is also capable of dealing with right-censoring and thus it is better suited to the analysis of recent agreements with ongoing ratifications.

Most of single-treaty survival studies focused the Kyoto Protocol (1997) and the UNFCCC (1992). Climate change agreements received a meticulous coverage not only because of their high media exposure, but also because of the rich anecdotal literature surrounding the manners negotiation was conducted and the debate behind participation in climate treaties. The COP [2] meetings are scrutinised by political scientist (Roger and Belliethathan 2016, Dimitrov 2016) and negotiation dynamics (Brandt and Svendsen 2004, Babiker et al. 2002), rules (Nasiritousi and Linner, 2016) and balances (Afionis 2011, Kasa et al. 2007) are carefully studied to explain countries' order of ratification (e.g.Andresen and Agrawala 2002, Lund 2013, Chin-Yee 2016). Survival analysis suits this branch of literature because it allows to test the ratification sequence in ways that are impossible with count data.

Neumayer (2002b), Wagner (2016) and Schneider and Urpelainen (2013) are the only papers that do not focus on climate agreements. The latter is an interesting study of the Cartagena Proctocol (2000), an agreement regulating the use of Living Modified Organisms (LMOs). The protocol puts forward the "precautionary principle" endorsed by the EU, which was thought to hinder the agricultural exports of United States by setting unfavourable international standards on LMOs. The United States strongly opposed the agreement and advocated the "sound science principle". Hence, the Cartagena Protocol is seen by the author as a natural experiment to test how political and diplomatic linkage with the Unites States and European Union affect the ratification behaviour of third states. Again, the choice of survival modelling is linked to the need of studying the sequence of ratification by different countries, which is easily performed with survival analysis.

2.2 Modelling choices

Among survival studies, the first difference in the methodological approach refers to the treatment of time. In many studies, time is treated as continuous even though models are based on yearly or monthly observations of ratification. Furthermore, the explanatory variables are always measured yearly. Hence, a common assumption is to take their values as constant throughout the year if the model is specified for monthly (von Stein 2008 and Schneider and Urpelainen 2013) or daily ratification (Fredriksson and Gaston 2000 and Fredriksson et al. 2007). The distinction between continuous and discrete observations is often a nuanced one. The ratification of an international agreement is *per se* a continuous process, however it is registered on time intervals of various length (years, months, weeks or days). Technically, it is a grouped survival data problem, because an underlying continuous process is observed discretely, hence the observations are grouped over an interval. So, despite the natural discreteness of the underlying data, depending on the granularity of the analysis, the variable could be assumed as continuous. Shorter observation intervals, such as days or weeks, over a long enough time period, could easily be considered a continuous representation of the ratification process. For annual observations the assumption is harder to justify (Neumayer, 2002a).

[2]Conference of the Parties (COP) is the annual meeting of the members of the UNFCCC (1992), Kyoto Protocol (1997) and the Paris Agreement (2015). National delegations gather to "keep under regular review the implementation of the Convention and any related legal instrument" (Art.7, UNFCCC 1992). COP meetings are attended by thousands of participants from NGOs, scientific organisations, universities, government bodies, industry representatives, media, and civil society in general.

Yamagata et al. (2013) and Sauquet (2014) are the only papers opting for a discrete approach.

In terms of model specification, the Cox proportional hazard model is the model of choice in the majority of the cases (Fredriksson and Gaston 2000, Neumayer 2002a, Fredriksson and Ujhelyi 2006, Fredriksson et al. 2007, von Stein 2008 and Schneider and Urpelainen 2013). Cox PH is a popular semi-parametric survival model that does not assume any particular distribution for the survival times. The shape of the baseline hazard remains unspecified, unlike in the Weibull and the Gompertz models used by Sauquet (2014). In proportional hazard models, the explanatory variables affect the hazard rate of ratification in a multiplicative fashion. Furthermore, the hazard ratio is assumed constant over time, implying that the relationship between the explanatory variable and the hazard ratio never changes. Proportional hazard models are different from accelerated failure time models which describe the speeding up process of an event. Wagner (2016) is the only ratification study that uses an accelerated failure time (AFT) model. In AFT models, the dependent variable is the ratification time instead of the hazard of ratification (probability of ratification at time t given no previous ratification). Except for Wagner (2016), all the models presented in this section are proportional hazard models and assume a baseline hazard shared among all the units of the analysis. It is a simplifying assumption that could clash with the structural diversity in ratification behaviours of nations. The samples contain diverse groups of nations but, except for Fredriksson and Ujhelyi (2006) and Fredriksson et al. (2007) that stratify their models on annex I and non-annex I countries, there has been no attempt to address unobserved heterogeneity at the country level.

Table 2: Survival analysis for single treaties

Paper	Sample	Dependent variable	Model
Fredriksson and Gaston (2000)	UNFCCC (1992), 184 countries until 1997.	Ratification survival time, daily observations.	Cox PH (also modelled as cross-sectional logistic regression).
Fredriksson et al. (2007)	Kyoto Protocol (1997), 170 countries until 2002.	Ratification survival time, daily observations.	Cox PH model stratified on annex I countries (also with a Weibull model and cross-sectional logistic regression).
von Stein (2008)	Kyoto Protocol (1997) and UNFCCC (1992), maximum of 140 countries until 2003.	Ratification survival time, monthly observations.	Separate models for the two treaties. Cox PH and Weibull specification.
Schneider and Urpelainen (2013)	Cartagena Proctocol (2000), 182 countries until 2006.	Ratification survival time, monthly observations.	Cox model allowing for non-proportional hazard. (also cross-sectional logit model).

Table 2: Survival analysis for single treaties (continued)

Paper	Sample	Dependent variable	Model
Yamagata et al. (2013)	Kyoto Protocol (1997) and UNFCCC (1992), maximum of 166 countries until 2008.	Ratification survival time, annual observations.	Logistic regression for discrete survival data with spatial lag (multiple weighting matrices used).
Sauquet (2014)	Kyoto Protocol (1997), 164 countries until 2009.	Ratification survival time, annual observations.	Gompertz survival model for grouped observations with spatial lag (weights based on trade, proximity and CDM projects).
Wagner (2016)	Montreal Protocol, Preferential Trade Agreements and Bilateral Investment treaties. 140 countries for the Montreal Protocol, until 2015.	Ratification survival time, daily observations.	Accelerated failure time model with spatial lag estimated with method of simulated moments (weights based on trade, IO membership and CFC emissions).

3 Pooled survival analysis

The survival approach can be extended to simultaneously deal with several treaties by pooling together the survival information of a group of treaties. Strictly speaking, the techniques used in this case are the same as in the previous section; the only difference is that, instead of dealing with countries, the unit of analysis is the country-treaty dyad. Bernauer et al. (2010) is the first study that pools together various treaties in a single survival model. Since then, this approach has been applied several times (see Table 4). Most of the recent studies choose to adopt this approach over the ones described in previous sections. This approach yields coefficient estimates that are general; they do not fit the specific treaty, instead they are intended to represent the process of ratification as a whole. From a methodological viewpoint, pooled survival models are more complex because they need to account also for the heterogeneity at the treaty level.

3.1 Ratification data sets

The first advantage of pooling different treaties together is that the number of observations is remarkably larger. The total size of the sample can be extended in any of the three dimensions of the analysis, by including more treaties, covering more countries or by lengthening the observation time. For each treaty-country dyad the beginning of the observation period corresponds to the signature year of the agreement and ends either with ratification by the country or on the cut-off year of the observation period. Most of the pooled survival studies use the ratification data collected by Bernauer et al. (2010)[3]. Their data set is notably larger than all previously used: It covers 180 countries and over 250 treaties. While earlier studies focused mainly on big environmental agreements, the data collected by Bernauer et al. (2010) allowed to diversify and expand the analysis to a profusion of smaller and lesser known agreements, considerably enriching the debate on ratification. In comparison, other data sets are relatively narrow in terms of countries and treaties. For example, Schulze (2014) only focuses on OECD countries and Leinaweaver (2012) cover 198 countries and only 55 agreements. Table 3 reports the sizes of a selection of large datasets used to study the ratification of environmental agreements.

Table 3: Ratification data sets

Data set	Treaties	Countries	Years	Regional treaties
Bellelli et al. (2020)	263	198	1950–2017	Yes
Bernauer et al. (2010)	255	180	1950–2000	No
Leinaweaver (2012)	55	193	1980–2010	Yes
Schulze and Tosun (2013)	21	25	1979–2010	Yes, all
Schulze (2014)	64	21	1971–2003	No
Cazals and Sauquet (2015)	41	99	1976–1999	No

The downside of pooling together many treaties is that it introduces the risk of sampling bias. In order to obtain generally valid ratification estimates, the sample needs not only to guarantee unbiasedness with respect to the mechanism of exclusion of countries from the sample, but also to be representative of the whole population of environmental treaties. Regrettably, in the context of previous studies, and except for Bellelli et al. (2020), the risks associated with sampling bias have not been thoroughly investigated and discussed. By construction, survival data on ratification has no discontinuity and is never left-censored, therefore missing observations occur among the explanatory variables rather than in the dependent. In larger studies we find no evaluation of the potential distortions deriving from the exclusion of countries with missing observations in the explanatory variables; and in the same way, the sensitivity of results to inclusion rules in the treaty sample has rarely been assessed. For example, regional environmental agreements have been either neglected or incorrectly handled in virtually all empirical studies. We now turn our attention to this category of treaties.

3.2 Mis-identification of potential ratifiers in regional treaties

Most of the ratification studies focus on *global* agreements (Bernauer et al. 2013a, Cazals and Sauquet 2015, Yamagata et al. 2017). These are those open to all nations and to which every nation is *de facto* a potential ratifier. Unfortunately, except for studies focusing on specific treaties or restricted to a group of countries (Perrin and Bernauer

[3]Their data set is used in the following works: Bernauer et al. (2010), Bernauer et al. (2013b), Bernauer et al. (2013a), Böhmelt et al. (2015), Mohrenberg et al. (2016), Spilker and Koubi (2016) and Hugh-Jones et al. (2018)

2010, Schulze and Tosun 2013, Schulze 2014, Yamagata et al. 2017), many less-than-global agreements have inadvertently been mixed with those with with global coverage. we call *regional* all the treaties that do not have strictly global coverage, without distinction for their scale or scope. The real concern is not so much that regional agreements have been included in the analysis, rather that they were incorrectly handled within the analysis. Note that most of the activities of environmental diplomacy take place at the regional level, therefore global agreement only represent a facet of international environmental cooperation (Leinaweaver, 2012). For regional agreements the situation is quite different: they have, by definition, a different set of potential ratifiers from those for global treaties. Unfortunately, in most of the literature it has always been implicitly assumed that all countries that did not ratify an agreement were either eligible or potentially capable of ratifying. As argued in Bellelli et al. (2020), this assumption holds for global treaties, but it becomes much less defensible when applied to regional agreements. In econometric terms it equates to incorrectly identifying the countries in the risk set. More specifically, it has been assumed that all existing countries are at risk of ratifying, while only a subset of them truly are. The resulting survival estimates are inevitably and systematically biased upward.

The data set assembled by Bernauer et al. (2010), and used in most of the studies, seems to be affected by this issue of mis-identification of potential ratifiers in regional agreements. There are good reasons to believe that a large fraction of their sample is indeed composed by regional agreements. Bernauer et al. (2010) are aware that some of the agreements could be *de facto* open only to a restricted number of countries. Hence, in their appendix they estimate a model exclusively on provenly global agreements, which results in their sample size being halved. Even in Leinaweaver (2012), where global and regional agreements are explicitly modelled jointly, the risk set appears to be incorrectly specified. Leinaweaver (2012) attempted to control for the regionality of a treaty by including dummies for the geographic regions of the ratifiers. However, this method is insufficient to correct the potential bias resulting from the erroneous specification of the risk set. The mis-identification of potential ratifiers was first exposed by Bellelli et al. (2020), who proposed an approach to correct the bias which consists in identifying the potential ratifiers for every environmental agreement in the sample.

Fortunately, the mis-identification bias we have just described does not affect all studies, as some studies with limited samples of either treaties or countries remain immune. For instance, Perrin and Bernauer (2010) and Schulze and Tosun (2013) exclusively focus on agreements negotiated under the UNECE[4]. Their analyses are confined to UNECE members because they perceive that non-UNECE nations may not ratify these treaties. With similar implications, Schulze (2014) exclusively focuses on the ratification by OECD nations, even if the agreements in their samples are open to other countries. Finally, Yamagata et al. (2017) is unaffected by the misidentification bias because their study is limited to eight agreements, all of which global.

3.3 Modelling choices and unobserved heterogeneity

In terms of choice of specification in the models, the studies differ mostly in two respects *i*) how time is defined and *ii*) in the manner unobserved heterogeneity is handled at

[4] United Nations Economic Commission for Europe

the treaty and country levels. Observations are always taken annually, except in Cazals and Sauquet (2015) who track ratification daily and assume the explanatory variables are constant over the year. A discrete treatment of time is prevalent. This approach involves expanding the survival data into a binary format in order to be explained by a binary regression model. Then, the baseline hazard is generally parameterised with either splines or cubic polynomials to allow for non-linearity. The estimates approximate those obtained with continuous survival models. The preferred modelling choice for continuous specifications of time is Cox PH models (Bernauer et al. 2013a, Schulze 2014, Cazals and Sauquet 2015 and Hugh-Jones et al. 2018).

With regards to unobserved heterogeneity, it can take place essentially at two levels: the country and the treaty level. We note, with some concern, that most of the studies with large samples (Bernauer et al. 2010, Perrin and Bernauer 2010, Böhmelt et al. 2015, Mohrenberg et al. 2016, Spilker and Koubi 2016) account for neither of these. Such shortcoming may justify some doubts on the consistency of the estimates. But, there are exceptions. For example, Cazals and Sauquet (2015) account for unobserved heterogeneity at the country level by including a "shared frailty" term in a continuous survival model. In survival analysis shared frailty is the equivalent of a country random effect[5]. Yamagata et al. (2017) control for treaty heterogeneity by including treaty dummies. Schulze (2014) and Hugh-Jones et al. (2018) account for heterogeneity across treaties by stratifying their models on the environmental subjects of the treaties (Hugh-Jones et al., 2018, see, for example,) or on each individual treaty (Schulze, 2014, as in). The problem with stratification is that it roughly corresponds to modelling each treaty (or group of treaties) separately. This type of solution rules out heterogeneity, but limits the ability to produce general inferences and it is harder to apply in large data sets. Finally, Leinaweaver (2012), Schulze and Tosun (2013) and Bellelli et al. (2020) are the only studies to date dealing with heterogeneity that can arise at both the country and the treaty level. These are modelled with random effects in a multilevel structure.

Table 4: Pooled survival analysis

Paper	Sample	Dependent variable	Model
Bernauer et al. (2010)	255 environmental agreements between 1950 and 2000, 180 countries.	Survival data on ratification recorded annually.	Binary regression for grouped survival data (also a cross-sectional logistic regression).

[5]Fixed effects are not usable with survival data because they perfectly predict non-occurrence. In other words, it would exclude all the units for which the event does not occur because their observations do not vary. The resulting survival estimates would be based solely on the units that experienced ratification.

Table 4: Pooled survival analysis (continued)

Paper	Sample	Dependent variable	Model
Perrin and Bernauer (2010)	9 Long-Range Transboundary Air Pollution (LRTAP) agreements, 47 Eurasian countries that ratified the 1979 convention. Between 1979 an 2007.	Survival data on ratification recorded annually.	Logistic regression for grouped survival data (also conditional logit with treaty fixed-effects).
Leinaweaver (2012)	55 environmental agreements (including regional) and 193 countries between 1980 and 2000.	Survival data on ratification recorded annually.	Logit model for survival data with country and treaty random effects.
Bernauer et al. (2013a)	286 agreements, 153 countries between 1973 and 2006.	Survival data on ratification recorded annually.	Cox PH model.
Schulze and Tosun (2013)	21 agreements negotiated under the UNECE. 25 non-EU countries between 1979 and 2010.	Survival data on ratification recorded annually.	Multilevel binary regression for discrete survival model with cross-classified random effects (Cox and logistic regression in appendix).
Cazals and Sauquet (2015)	41 environmental agreements ratification by 99 countries from 1976 to 1999.	Survival data on ratification recorded daily.	Cox PH model with frailty term shared at country level.
Böhmelt et al. (2015)	250 agreements, 75 democracies between 1973 and 2002. Data from Bernauer et al. (2010).	Survival data on ratification recorded annually.	Logistic regression for survival data.
Mohrenberg et al. (2016)	219 agreements, 160 countries between 1950 and 2000. Data from Bernauer et al. (2010).	Survival data on ratification recorded annually.	Logistic regression for survival data.

Table 4: Pooled survival analysis (continued)

Paper	Sample	Dependent variable	Model
Spilker and Koubi (2016)	220 agreements, 162 countries between 1950 and 2000. Data from Bernauer et al. (2010).	Survival data on ratification recorded annually.	Logistic regression for survival data.
Yamagata et al. (2017)	8 agreements and 166 countries between 1981 and 2006.	Survival data on ratification recorded annually.	Two separate logit regressions (pre- and post-1991) for discrete survival analysis. Spatial lag with multiple weighting matrices.
Hugh-Jones et al. (2018)	126 agreements and 157 countries between 1972 and 2000. Bernauer et al. (2010).	Survival data on ratification recorded annually.	Cox PH model stratified on different areas of regulation.
Bellelli et al. (2020)	258 agreements and 192 countries between 1990 and 2015.	Survival data on ratification recorded annually	Cross-classified multi-level discrete survival model with country and treaty random effects estimated with Markow Chain Monte Carlo (MCMC).

4 Concluding remarks

In this appendix, we described the evolution of the three methodological approaches used to empirically study the ratification of environmental agreements. At first, empirical studies use mostly a cross-sectional approach with ratification count data. This approach has numerous limitations, for instance, estimates may be influenced by the cut-off date. Count data models do not allow to explore at the same time country's and treaty's characteristics. Also, the total number of ratification is an opaque measure: does participation in more treaties by a country really imply stronger environmental commitments? The number of ratified agreements largely depends on the number of agreements the country can potentially ratify, a factor that has never been accounted for in this type of studies. The fundamental problem of count data is that it does not allow to identify how countries differ in their ratification choices for the same agreement. Finally, this approach does not cast any light on how ratification by a country interacts with decisions by other countries because these approaches do not use information regarding the timing and hence the sequence of ratification decisions.

Given these limitations, the methodological approach gradually shifted towards the use of survival models. These allow researchers to study both the occurrence and the timing of ratification. We found it useful to distinguish between studies focusing on single agreements and those studying a pooled sample of treaties. Analyses based on the survival approach tackle most of the shortcomings of the previous methodology: it can easily cope with right-censoring, ratifications can be traced to the treaty and country (therefore treaty and country variables can be studied jointly). And importantly, survival analysis allows to study the differences in ratification timing, consequently researcher can study how ratification decisions by different countries interact with each other. However, we note that survival models for large samples of agreements face methodological complexities which are not always appropriately tackled, or even adequately discussed. Above all, pooled survival models need to address the unobserved heterogeneity at the treaty and country level, and ensure the correct identification of potential ratifiers. Unfortunately, we found that in most of the empirical literature to date these issues have been inadequately addressed.

References

Afionis, S. (2011). The European Union as a negotiator in the international climate change regime. *International Environmental Agreements: Politics, Law and Economics*, 11(4):341–360.

Almer, C. and Winkler, R. (2010). Strategic behavior in IEAs: When and why countries joined the Kyoto Protocol. *Bern University Discussion Papers*, 14:1–28. (Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1549612).

Andresen, S. and Agrawala, S. (2002). Leaders, pushers and laggards in the making of the climate regime. *Global Environmental Change*, 12:41–51.

Babiker, M. H., Jacoby, H. D., Reilly, J. M., and Reiner, D. M. (2002). The evolution of a climate regime: Kyoto to Marrakech and beyond. *Environmental Science & Policy*, 5:195–206.

Bellelli, F. S., Scarpa, R., and Aftab, A. (2020). Why do countries join international environmental agreements? A look at the influence of interest groups. *Working paper.*

Bernauer, T., Böhmelt, T., and Koubi, V. (2013a). Is there a democracy-civil society paradox in global environmental governance. *Global Environmental Politics*, 13(1):88–107.

Bernauer, T., Kalbhenn, A., Koubi, V., and Spilker, G. (2010). A comparison of international and domestic sources of global governance dynamics. *British Journal of Political Science*, 40:509–538.

Bernauer, T., Kalbhenn, A., Koubi, V., and Spilker, G. (2013b). Is there a "Depth versus Participation" dilemma in international cooperation. *The Review of International Organizations*, 8:477–497.

Beron, K., Murdoch, J., and Vijverberg, W. (2003). Why cooperate? Public goods, economic power, and the Montreal Protocol. *The Review of Economics and Statistics*, 85(2):286–297.

Biodiversity Convention (1992). Convention on biological diversity. *Secretariat of the Convention on Biological Diversity*. Rio de Janeiro. (available at: https://www.cbd.int/convention/text/).

Böhmelt, T., Bernauer, T., and Koubi, V. (2015). The marginal impact of ENGO in different types of democratic systems. *European Political Science Review*, 7(1):93–118.

Brandt, U. S. and Svendsen, G. T. (2004). Fighting windmills: The coalition of industrialists and environmentalists in the climate change issue. *International Environmental Agreements: Politics, Law and Economics*, 4:327–337.

Cartagena Proctocol (2000). Cartagena protocol on biosafety to the convention on biological diversity. *Secretariat of the Convention on Biological Diversity*. (available at: https://bch.cbd.int/protocol/text/).

Cazals, A. and Sauquet, A. (2015). How do elections affect international cooperation? Evidence from environmental treaty participation. *Public Choice*, 162:263–285.

Chin-Yee, S. (2016). Briefing: Africa and the Paris climate change agreement. *African Affairs*, 115(459):359–268.

CITES (1973). Convention on international trade in endangered species of wild fauna and flora. *CITES Secretariat*. Washington. (Available at: https://www.cites.org/eng/disc/text.php).

Congleton, R. (1992). Political institutions and pollution control. *The Review of Economics and Statistics*, 74(3):412–421.

Davies, R. and Naughton, H. (2014). Cooperation in environmental policy: a spatial approach. *International Tax and Public Finance*, 21(5):923–954.

Dimitrov, R. (2016). The Paris Agreement on climate change: Behind closed doors. *Global Environmental Politics*, 16(3):1–11.

Egger, P., Jeßberger, C., and Larch, M. (2011). Trade and investment liberalization as determinants of multilateral environmental agreement membership. *International Tax and Public Finance*, 18(6):605–633.

Egger, P., Jessberger, C., and Larch, M. (2013). Impacts of trade and the environment on clustered multilateral environmental agreements. *The World Economy*, 36(3):331–348.

Frank, D. J. (1999). The social bases of environmental treaty ratification. *Sociological Inquiry*, 69(4):523–550.

Fredriksson, P. and Gaston, N. (2000). Ratification of the 1992 climate change convention: What determines legislative delay? *Public Choice*, 104:345–368.

Fredriksson, P., Neumayer, E., and Ujhelyi, G. (2007). Kyoto Protocol cooperation: Does government corruption facilitate environmental lobbying? *Public Choice*, 133:231–251.

Fredriksson, P. and Ujhelyi, G. (2006). Political institutions, interest groups, and the ratification of international environmental agreements. *Working paper*.

Helsinki Protocol (1985). Helsinki Protocol on the Reduction of Sulphur Emissions. *UNECE Secretariat*. (Available at: https://www.unece.org/env/lrtap/sulf_h1.html).

Hugh-Jones, D., Milewicz, K., and Ward, H. (2018). Signaling by signature: The weight of international opinion and ratification of treaties by domestic veto players. *Political Science Research and Methods*, 6(1):15–31.

Kasa, S., Gullberg, A., and Heggelund, G. (2007). The group of 77 in the international climate negotiations: recent developments and future decisions. *International Environmental Agreements: Politics, Law and Economics*, 8:113–127.

Kyoto Protocol (1997). Kyoto Protocol to the United Nations Framework Convention on Climate Change. *UNFCCC Secretariat*. Kyoto. (Available at: http://unfccc.int/kyoto_protocol/items/2830.php).

Leinaweaver, J. (2012). Environmental treaty ratification: Treaty design, domestic politics and international incentives. *Working paper*. (available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2035531).

Lund, E. (2013). Environmental diplomacy: Comparing the influence of business and environmental NGOs in negotiations on reform of the Clean Development Mechanism. *Environmental Politics*, 22(5):739–759.

Mohrenberg, S., Koubi, V., and Bernauer, T. (2016). Ratification of multilateral environmental agreements: Civil society access and financial mechanisms. *National Centre of Competence Research (NCCR)*, Working Paper No.91. (Available at: http://www.nccr-democracy.uzh.ch/publications/workingpaper/pdf/wp_91.pdf).

Montreal Protocol (1987). The Montreal Protocol on Substances that Deplete the Ozone Layer. *UNEP Ozone Secretariat*. Montreal. (Available at: http://ozone.unep.org/en/treaties-anddecisions/montreal-protocol-substances-deplete-ozone-layer).

Murdoch, J., Sandler, T., and Vijverberg, W. (2003). The participation decision versus the level of participation in an environmental treaty: A spatial probit analysis. *Journal of Public Economics*, 87:337–363.

Nasiritousi, N. and Linner, B.-O. (2016). Open or closed meetings? Explaining non-state actor involvement in the international climate change negotiations. *International Environmental Agreements: Politics, Law and Economics*, 16(1):127–144.

Neumayer, E. (2002a). Do democracies exhibit stronger environmental commitment? A cross-country analysis. *Journal of Peace Research*, 39(2):139–164.

Neumayer, E. (2002b). Does trade openness promote multilateral environmental cooperation? *World Economy*, 25(6):815–832.

Paris Agreement (2015). The Paris Agreement. *UNFCCC Secretariat*. Paris. (Available at: http://unfccc.int/paris_agreement/items/9444.php).

Perrin, S. and Bernauer, T. (2010). International regime formation revisited: Explaining ratification behaviour with respect to long-range transboundary air pollution agreements in Europe. *European Union Politics*, 11(3):405–426.

Recchia, S. (2002). International environmental treaty engagement in 19 democracies. *Policy Studies Journal*, 30(4):470–494.

Roberts, J., Parks, B., and Vasquez, A. (2004). Who ratifies environmental treaties and why? Institutionalism, structuralism and participation by 192 nations in 22 treaties. *Global Environmental Politics*, 4(3):22–64.

Roberts, J. T. (1996). Predicting participation in environmental treaties: A world-system analysis. *Sociology Inquiry*, 66(1).

Roger, C. and Belliethathan, S. (2016). Africa in the global climate change negotiations. *International Environmental Agreements: Politics, Law and Economics*, 16:91–108.

Sauquet, A. (2014). Exploring the nature of inter-country interactions in the process of ratifying international environmental agreements: the case of the Kyoto Protocol. *Public Choice*, 159:141–158.

Schneider, C. and Urpelainen, J. (2013). Distributional conflict between powerful states and international treaty ratification. *International Studies Quarterly*, 57:13–27.

Schulze, K. (2014). Do parties matter for international environmental cooperation? An analysis of environmental treaty participation by advanced industrialised democracies. *Environmental Politics*, 23(1):115–139.

Schulze, K. and Tosun, J. (2013). External dimensions of European environmental policy: An analysis of environmental treaty ratification by third states. *European Journal of Political Research*, 52:581–607.

Seelarbokus, C. (2014). Theorizing state participation in international environmental agreements (IEAs). *World Environment*, 4(3):121–142.

Spilker, G. and Koubi, V. (2016). The effects of treaty legality and domestic institutional hurdles on environmental treaty ratification. *International Environmental Agreements: Politics, Law and Economics*, 16:223–238.

UNFCCC (1992). United Nations Framework Convention on Climate Change. *UNFCCC Secretariat*. Rio de Janeiro. (Available at: http://unfccc.int/essential_background/convention/items/6036.php).

Vienna Convention (1985). Vienna Convention on the Protection of the Ozone Layer. *UNEP Ozone Secretariat*. Vienna. (Available at: http://ozone.unep.org/en/treaties-and-decisions/vienna-conventionprotection-ozone-layer).

von Stein, J. (2008). The international law and politics of climate change. *Journal of Conflict Resolution*, 52(2):243–268.

Wagner, U. (2016). Estimating strategic models of international treaty formation. *Review of Economic Studies*, 83(4):1741–1778.

Yamagata, Y., Yang, J., and Galaskiewicz, J. (2013). A contingency theory of policy innovation: How different theories explain the ratification of the UNFCCC and the Kyoto Protocol. *International Environmental Agreements: Politics, Law and Economics*, 13:251–270.

Yamagata, Y., Yang, J., and Galaskiewicz, J. (2017). State power and diffusion process in the ratification of global environmental treaties, 1981-2008. *International Environmental Agreements: Politics, Law and Economics*, 17(4):501–529.