Review of Economic Studies (2016) **83**, 1741–1778 doi:10.1093/restud/rdv054

Advance access publication 20 February 2016

Estimating Strategic Models of International Treaty Formation

ULRICH J. WAGNER

University of Mannheim

First version received July 2011*; final version accepted July* 2015 (*Eds.*)

This article develops an empirical framework for analysing the timing of international treaties. A treaty is modelled as a dynamic game among governments that decide on participation in every period. The net benefit of treaty membership increases over time. Spillovers among members and non-members accelerate or delay treaty formation by transforming participation into a strategic complement or substitute, respectively. The predictions of the model inform the estimation of the structural parameters, based on a cross section of treaty ratification dates. With this approach, I estimate the sign and magnitude of strategic interaction in the ratification of the Montreal Protocol, in the formation of Europe's preferential trade agreements, and in the growth of Germany's network of bilateral investment treaties. Through a series of counterfactual experiments, I explore different mechanisms that give rise to strategic interaction in the formation of these treaties.

Key words: Dynamic game, strategic interaction, structural estimation, International environmental agreement, preferential trade agreement, bilateral investment treaty, international relations.

JEL Codes: C51, F53, F18, Q58, H41, K33

1. INTRODUCTION

What drives countries to cooperate in the pursuit of common policy goals? Ever since World War II, cross-border policy coordination has been on the rise in many areas, including military defence, human rights, international trade and finance, public health, and environmental protection. While the specific reasons behind such alliances differ widely, as a rule they materialize in situations where coordinated efforts are better suited to achieve the common good than unilateral policies. This suggests that collective rationality is a driving force behind international cooperation. However, there are numerous international policy issues to which collective rationality would dictate a cooperative approach and yet international cooperation fails badly. This failure to provide global public goods or to manage common property resources efficiently is often attributed to the adverse incentives faced by individual governments. If countries do not take into account the external benefits of their contribution to the public good, the aggregate provision level falls short of the social optimum (Samuelson, 1954). Worse, even though the provision of the public good is desirable for all countries, some of them may prefer to free-ride by enjoying the external benefits without sharing the cost.

At the national level, such conflicts between individual and collective rationality can be resolved by the intervention of the government (Demsetz, 1967). At the international scale, however, there is no supranational authority that could coerce states into adopting efficient policies if they run counter to national interests. Filling the void are international agreements. Under the terms of the Vienna Convention on the Law of Treaties, a state that ratifies a multilateral treaty chooses partially to surrender its sovereignty and to subject its policies in a specific domain to the rules and prescriptions of the treaty.[1] In so doing, sovereign states agree to coordinate their policies in mutually beneficial ways. By the very nature of sovereignty, however, the agreement is fundamentally non-binding and states can always withdraw from it. Therefore, the fact that public good provision is implemented through an international agreement should not change a country's incentives to contribute per se—unless the treaty alters the country's incentives to cooperate in other ways.

This article proposes a new empirical framework to shed light on these incentives. I develop a structural econometric model that exploits the variation in ratification dates to estimate the sign and magnitude of strategic interaction. A treaty is modelled as a repeated game among governments that choose between two actions in every period: to participate or not. The difference in payoffs to both actions—the relative payoff to cooperation—is assumed to increase monotonically over time. This accommodates the possibility that a country's attitude towards treaty participation changes for reasons unrelated to other countries' actions, such as an exogenous decline in compliance cost or an increasing valuation of the treaty's benefits.

My approach nests this hypothesis into a model where interdependent ratification behaviour may arise as a consequence of strategic choices by forward-looking governments. To allow for rich patterns of strategic interaction, I consider three types of spillovers. First, treaty membership may generate positive spillovers to everyone else, reflecting the fact that public good provision is the key objective in a large class of international agreements. Secondly, this spillover may differ between treaty participants and non participants. Consequently, the relative payoff to cooperation can either decrease or increase with the number of treaty members. If ratification becomes less desirable as treaty membership grows, it is a strategic substitute (Bulow *et al.*, 1985). This parallels the free-riding incentive in a public good game with decreasing marginal benefits. In contrast, if the relative benefit to cooperation increases with the number of treaty members, ratification is a strategic complement. Strategic complementarity may arise for a variety of reasons, including social preferences or concerns about reputation held by governments, discreteness of the public good the treaty seeks to provide, or aspects of the treaty design itself. Thirdly, when ratification is a strategic complement, the ratification game is supermodular and hence easy to solve even when spillovers vary across country pairs. I exploit this feature to empirically test specific channels through which strategic complementarity affects treaty formation.

The assumptions on the payoff function strike a balance between the need to accommodate complex dynamic interactions arising in a broad class of treaties and the necessity of keeping the empirical model simple enough to be viable in settings where data availability is often limited to a single ratification history. The monotonicity assumption ensures that all countries join the treaty eventually, which allows me to solve the dynamic game via backward induction. In the payoff-maximizing subgame-perfect equilibrium, strategic substitutability delays treaty ratification whereas complementarity has the opposite effect. If strategic complementarity is sufficiently strong, ratification by one country triggers ratification by another, generating clusters of ratification events in the data. These predictions identify the sign and magnitude of strategic interaction in a structural econometric model of treaty ratification. Moreover, the structural model

1. The legal procedure of a state joining a multilateral agreement is the signature followed by ratification, which marks the legal accession. In this study, I will be using the terms "join", "accede", "ratify", "participate" interchangeably, referring to the legal act of accession as opposed to a mere signing of a treaty.

allows me to run counterfactual experiments to explore the role of such channels with a view to optimal treaty design.

I apply this method to the Montreal Protocol on Substances that Deplete the Ozone Layer which was opened for signature in 1987 and has since been ratified by 191 countries. Its Member States have committed to phasing out emissions of chlorofluorocarbons (CFCs) and other pollutants that, via the depletion of stratospheric ozone, cause severe damages to human health and to the natural environment. The Montreal Protocol is credited with accomplishing the provision of a pure public good at the global scale, overcoming countries' incentive to free-ride.

The success of the Montreal Protocol has been studied widely, not least because it seemed to offer important lessons for the 1997 Kyoto Protocol whose targets and timetables for greenhouse gas abatement closely followed the design of the Montreal Protocol. Earlier analyses attribute much of the success of the Montreal Protocol to participation being a dominant strategy for the dominant producer—the U.S.—and to the trade ban which deterred free-riding by fringe producers (Benedick, 1998; Barrett, 2003, 1997b; Parson, 2003; Sunstein, 2007). This literature reveals that incentives to participate in the Montreal Protocol depend on a complex combination of factors, including the economics of ozone depletion, the industrial organization of the polluting industries, specific features of the treaty design and aspects of international diplomacy more broadly. While this literature has followed a qualitative approach, I offer a strictly quantitative perspective on this important policy issue. Moreover, taking a model-based approach to estimation allows me to tease apart the moving parts of the mechanisms that drive participation through a series of counterfactual explorations.

In further applications, I demonstrate how my framework can be employed to study strategic interaction in the diffusion of international treaties on trade and investment. It has been suggested that the rapid proliferation of such treaties over the last 60 years can be explained in part by strategic behaviour which arises because the treaties impose external costs on third countries. For instance, the diversion of trade flows caused by a regional trade agreement may lead to a "domino effect" by which non-member countries also slash tariffs and other trade barriers (Baldwin, 1993). Likewise, a country can see its prospects for receiving foreign direct investment diminished when a rival country signs a bilateral investment agreement that grants enhanced protection to the foreign investor. While these theories suggest that participation is a strategic complement, there is no conclusive evidence yet as to whether this matters empirically. To shed light on this, I apply my framework to data on Europe's preferential trade agreements (PTAs) and on the formation of Germany's network of bilateral investment treaties (BITs). This allows me to empirically test whether the mechanisms proposed in the literature actually did create strategic complementarities in the formation of these treaties.

Relation to the literature This article extends the empirical literature on timing games in oligopolistic settings (*e.g.* Schmidt-Dengler, 2006; Sweeting, 2006; Einav, 2010) by solving and estimating a discrete dynamic game of complete information where actions can be either strategic substitutes or complements. Similar to independent work by de Paula (2009) it is shown that clustering occurs as a consequence of strategic complementarity in the payoff functions.[2] His contribution and the analysis in this article are complementary in the sense that they derive this result starting with fundamentally different assumptions about the information structure

2. In a "synchronization game" with social interactions, de Paula (2009) derives a test for strategic complementarity based on simultaneous stopping by multiple agents. Ackerberg and Gowrisankaran (2006) also exploit this for identification, even though not in the context of a timing game.

(incomplete versus complete), the time structure (continuous versus discrete), the strategy space (irreversible versus reversible actions), and the solution concept.

From an econometric point-of-view, the framework presented in this article is an alternative to duration models which Honoré and de Paula (2010) only recently extended to account for strategic behaviour in a game with two players and irreversible actions. This article adds to their work by developing an N-player game with reversible actions, which is a common setting also in the literature on dynamic oligopoly (Aguirregabiria and Mira, 2010). The methods developed in that literature require a data set sufficiently rich to estimate state transition probabilities and policy functions for all players, which is impractical in the context of a single treaty. The advantage of the method proposed in this article lies in the fact that it is feasible in settings with limited data availability, including the case where only a single history of the game is observed. This is achieved by limiting the set of payoff-relevant state variables to a time effect.

The application to the Montreal Protocol contributes to a sizeable literature on international environmental agreements (see Wagner, 2001, for a survey). The workhorse in this literature has been the "self-enforcing" agreements model (Hoel, 1992; Carraro and Siniscalco, 1993; Barrett, 1994) which models treaty participation as a non-cooperative game played among governments. While the basic version of this model is static, some authors have also considered a supergame to allow for a richer strategies space (*e.g.* Barrett, 1994; Finus and Rundshagen, 1998; Asheim and Holtsmark, 2009) or the differential game to allow for richer dynamics (Rubio and Ulph, 2007; de Zeeuw, 2008). This article develops and estimates the first model to explain the *timing* of ratification across countries. The model accommodates heterogeneous payoff functions with asymmetric strategic interactions—features that engender multiple equilibria and may prevent an analytical solution in the self-enforcing agreements model (Barrett, 1997a; McGinty, 2007). Since that model's key prediction is the number of treaty members, its empirical implementation is complicated given the lack of a sample of independent and identical treaties. Consequently, previous empirical work has taken reduced-form approaches (Murdoch and Sandler, 1997; Beron *et al.*, 2003; Finus and Tjøtta, 2003; Murdoch *et al.*, 2003; Bratberg *et al.*, 2005; Wagner, 2009; Bernauer *et al.*, 2010; Aakvik and Tjøtta, 2011), limited the number of agents to two stylized players (Auffhammer *et al.*, 2005) or ignored strategic interaction altogether (Congleton, 1992; Fredriksson and Gaston, 2000; Neumayer, 2002). This article bridges the gap between theory and empirics by developing a strategic model of treaty formation which is estimable using a cross-section of ratification times. Since information is complete, countries condition their actions on observed participation rather than on unobserved latent benefits as in Beron *et al.* (2003). Furthermore, since endogenous effects are identified off the time intervals between subsequent ratification decisions, *a priori* there is no need to specify a political or economic channel through which interactions work—though the relevance of such channels can and will be tested.

By applying my framework to PTAs, I contribute new evidence pertaining to the long-standing controversy over whether regional trade integration is a stepping stone or a stumbling block for multilateral trade liberalization (Baldwin, 1993; Panagariya, 1999; Bhagwati, 2008). The application to BITs contributes new evidence on the endogenous formation of institutions that govern foreign direct investment. Although the significance of such institutions for the efficiency of international capital mobility can hardly be overstated, BITs have received little attention by economists so far. There is, however, a sizeable literature on this topic in the field of international relations. A more detailed discussion of how the present article extends the empirical literatures on PTAs and BITs is relegated to Section 5 below.

The remainder of the article is structured as follows. Section 2 develops a theoretical model of the timing of treaty ratification and derives the key testable implications. Section 3 explains how these implications are exploited for identification and estimation of the model parameters.

Section 4 applies the framework to study the ratification of the Montreal Protocol and Section 5 discusses further applications to trade and investment agreements. Section 6 concludes.

2. A TIMING GAME OF TREATY PARTICIPATION

The goal of this article is to develop an empirical framework that allows researchers (1) to infer the nature and channels of strategic interaction from a history of participation decisions, and (2) to conduct counterfactual experiments. The structural approach is based on a game theoretical model that should be general enough to fit a broad class of treaties, while also being specific enough to generate testable predictions on equilibrium play that inform the estimation of the model parameters. This section introduces the model and explains how these requirements condition the modelling choices.

2.1. *Model setup*

Treaty participation is modelled as an infinite-horizon game played among N countries. Countries are assumed to have complete information about all primitives of the game. In each period $t \in T = \{0,1,\dots\}$ all countries choose simultaneously whether to participate in the treaty ($a_i = 1$) or not ($a_i = 0$).[3] In line with the Vienna Convention on the Law of Treaties, countries can costlessly withdraw from the treaty. Country i's continuation payoff in period t under strategy profile $s \in S_t$ is given by

$$V_i(s,t) = \sum_{\tau=t}^{\infty} \pi_i(a_i(s,\tau), a_{-i}(s,\tau), \tau) e^{-r_i(\tau - t)} \tag{2.1}$$

where $\pi_i(\cdot)$ denotes the per-period payoff function of country i, $a(s,t)$ is the action profile induced at time t by strategy profile s and $r_i > 0$ is country i's discount rate. Each country i chooses its strategy s_i so as to maximize $V_i(s_i, s_{-i}, 0)$ taking as given the strategies s_{-i} chosen by the other countries. The following transversality condition ensures that this programme has a well-defined solution as time goes to infinity

Assumption 1. $\lim_{t\to\infty} e^{-r_i t} \pi_i(a,t) = 0 \quad \forall i \in I \ \forall a \in A.$

Assume further that the per-period payoff takes the form

$$\pi_i(a_i(t), a_{-i}(t), t) = \begin{cases} \gamma_0 \sum_{j \neq i} w_{ij} a_j(t) & \text{if } a_i = 0 \\ -\phi_i + g(t) + \gamma_1 \sum_{j \neq i} w_{ij} a_j(t) & \text{if } a_i = 1, \end{cases} \tag{2.2}$$

where $\gamma_0, \gamma_1, \{\phi_i\}_{i \in I}$ are parameters and the weights $w_{ij} \in [0,1] \forall i,j \in I$ measure the intensity of country j's participation decision on country i's payoff.

This specification admits the possibility of both exogenous and endogenous changes in the incentive to participate. This incentive is characterized by the *relative payoff to cooperation*, parameterized as

$$\Delta\pi_i(a_{-i},t) \equiv \pi_i(1,a_{-i},t) - \pi_i(0,a_{-i},t) = -\phi_i + g(t) + \gamma \sum_{j \neq i} w_{ij} a_j(t) \tag{2.3}$$

where $\gamma \equiv \gamma_1 - \gamma_0$. The sum of the first two terms determines a country's private incentive to ratify the treaty.

3. I adopt standard notation, *i.e.* I denotes the set of countries, $A_i = \{0,1\}$ the action set, and $A = \prod_{i \in I} A_i$ the action space with generic element $a = (a_i, a_{-i})$, where $a_{-i} \in A_{-i} \equiv \prod_{j \neq i} A_j$ is the vector of actions taken by all countries other than i.

The following assumptions on the payoff function are designed to pin down three common features that determine the level of participation in a broad set of international treaties. First, I allow for positive externalities that arise in treaties that provide a public good.

Assumption 2. *$\forall i \in I$, $\forall t \in T$, $\forall a, a' \in A$ such that $a_{-i} \le a'_{-i}$, $a_{-i} \ne a'_{-i}$;*

$$\pi_i(a_i, a_{-i}, t) < \pi_i(a_i, a'_{-i}, t).$$

Next, to allow for an exogenous driver of treaty participation I assume that the relative payoff to cooperation increases over time, *i.e.* $\Delta\pi_i(a_{-i}, t+1) > \Delta\pi_i(a_{-i}, t) \forall a_{-i} \in A_{-i}$.

Assumption 3. *The function $g(t)$ is strictly increasing and differentiable in calendar time.*

This assumption nests the hypothesis of non-strategic ratification behaviour. If $\gamma = 0$ then the incentive to cooperate is invariant with respect to the number and identity of other contributors, *i.e.* $\Delta\pi_i(a_{-i}, t) = \Delta\pi_i(a'_{-i}, t) \forall a_{-i}, a'_{-i} \in A_{-i}$.

Alternatively, the relative payoff to cooperation may increase or decrease if an additional country joins the treaty. The sign and magnitude of this strategic effect depend on the parameter γ. If $\gamma > 0$ then the relative payoff to cooperation exhibits increasing differences $\Delta\pi_i(a_{-i}, t) \le \Delta\pi_i(a'_{-i}, t) \quad a_{-i} \le a'_{-i}, a_{-i} \ne a'_{-i}$ and participation in the treaty is a strategic complement as long as the inequality is strict for at least one country. In turn, if $\gamma < 0$ the relative payoff to participation decreases when an additional country joins the treaty and hence participation is a strategic substitute.

This simple yet general specification encompasses a broad set of models and settings with rich patterns of strategic interaction that may affect the timing of treaty ratification. For instance, in the case of a pure public good one would set $w_{ij} = b$ to reflect non-excludable external benefits and assume constant ($\gamma_0 = \gamma_1$) or decreasing ($\gamma_1 < \gamma_0$) marginal benefits of contributing to the public good. In contrast, if the treaty bans Member States from trading with non-Member States, then the incentive to participate increases as a country's most important trading partners join the treaty.[4] To reflect this, one would calibrate w_{ij} on bilateral exports and assume $\gamma_1 > \gamma_0$.

Notice that strategic complementarity or substitutability can also arise in more sophisticated models of public good provision. In a differential game setting, contributions to a public good with decreasing marginal benefits are strategic substitutes when the public good is continuous (Fershtman and Nitzan, 1991) and complements when the public good is discrete (Kessing, 2007; Georgiadis, 2015). The simple framework proposed here gives up some of the rigor and elegance of those theoretical models in order to give empirical tractability to the estimation of strategic interaction based on readily available ratification data. As I will show next, the one-parameter specification of strategic interaction is sufficient to generate sharp predictions on equilibrium play. These predictions allow me to estimate fundamental parameters of the model from the data, without assuming the nature of strategy interaction a *priori*.

2.2. *A graphical illustration of equilibrium with two countries*

To gain some intuition for the solution of the dynamic game, consider the graphical representation of a simple two-country version of the game in Figure 1. Figure 1(a) depicts the time path of the relative payoff to cooperation (2.3) for both countries under the assumption that $\gamma = 0$ (constant

4. I explain this mechanism in more detail in Section 4 below, and in Section C of the Online Appendix.

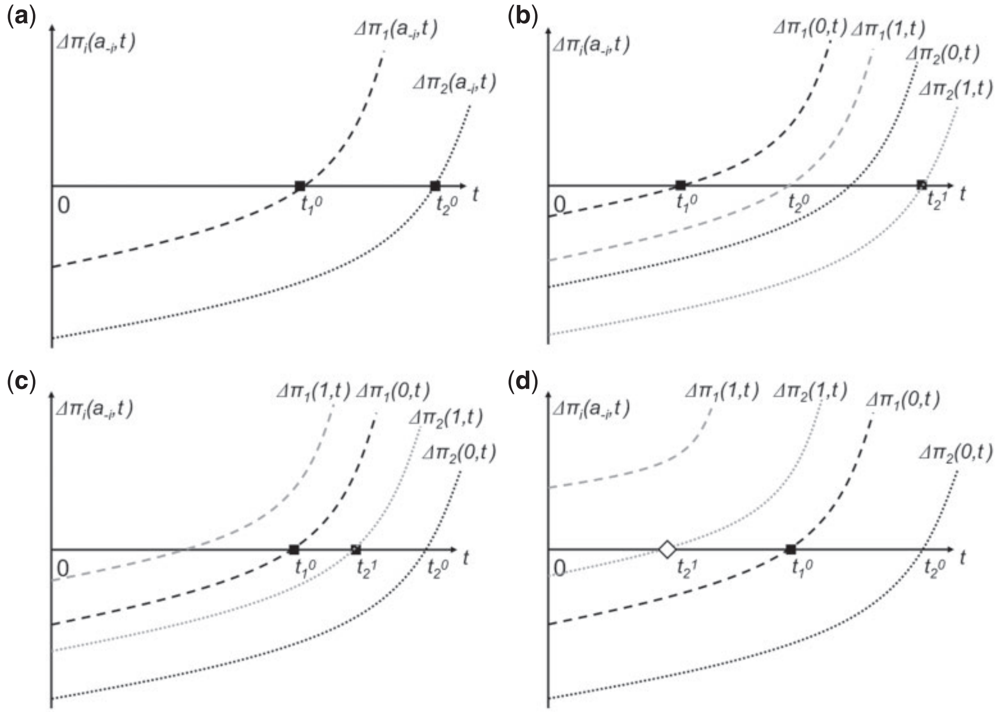

FIGURE 1

Evolution of stage-game payoffs over time. (a) Constant differences. (b) Strategic substitutability. (c) Weak strategic complementarity. (d) Strong strategic complementarity.

differences). It is easily seen that country i's dominant strategy is to join the treaty only from period t_i^0 onwards.

The case of strategic substitutability ($\gamma < 0$) is depicted in Figure 1(b). While the first country joins the treaty at the same time as in Figure 1(a), the negative payoff externality means that the relative payoff to participation shifts downwards for country 2 and hence delays ratification of the treaty from period t_2^0 until t_2^1.

In contrast, if ratification is a strategic complement ($\gamma > 0$), Figure 1(c) shows how the relative payoff curve of country 2 shifts upward once country 1 joins the treaty in period t_1^0, so that cooperation is country 2's best response from period t_2^1 onwards. Hence strategic complementarity accelerates treaty formation compared to the base case depicted in Figure 1(a). With strong strategic complementarity ($\gamma \gg 0$), the payoffs can be depicted as in Figure 1(d). Both countries start out as non signatories but now both of them would be better off as treaty members even before period t_1^0 is reached. In fact, both cooperation and no cooperation are Nash equilibria at every stage in the interval $[t_2^1, t_1^0)$. Because of positive externalities (Assumption 2), both countries prefer the cooperative Nash equilibrium to the non-cooperative one. Below I show that the equilibrium refinement of strong renegotiation-proof equilibrium selects a unique subgame-perfect equilibrium in which both players start to cooperate at the earliest possible period t_2^1 and never revert to non-cooperative behaviour. An important implication of the case $\gamma > 0$ is that even asymmetric countries may begin to cooperate in the same period, as part of a cluster of cooperation decisions. The empirical strategy adopted below exploits all these strategic effects for parameter identification, *i.e.* the delayed versus accelerated ratification by late adopters depicted in Figure 1(b) and (c) as well as the feature of clustered participation decisions depicted in Figure 1(d).

2.3. *Equilibrium*

This section analyses the equilibrium of the dynamic game under different assumptions about the sign of γ. I start by characterizing the pure-strategy Nash equilibria of the stage-game $\Gamma(t) = (I, A, \pi(t))$ played in period t, where $\pi(t)$ is a vector of individual payoff functions $\pi_i(t)$ for all countries. I then characterize the subgame-perfect equilibria of the infinite-horizon game $G = \{I, S, V(0)\}$, formed by the infinite sequence of stage games $\Gamma(t)_{t=0}^{\infty}$. All proofs are relegated to the Appendix.

For the case of $\gamma \geq 0$, the theory of supermodular games has established that pure-strategy Nash equilibria of the stage game exist, that they are Pareto-ranked and monotonic in time.

Theorem 1. *If $\gamma \geq 0$, the stage game $\Gamma(t)$ has a smallest and a largest (in terms of the number of treaty members) Nash equilibrium in pure strategies for all $t \in T$. Both Nash equilibria are weakly increasing in t. All players prefer the largest Nash equilibrium over any other Nash equilibrium.*

Corollary 1. *If $\gamma = 0$ the pure-strategy Nash equilibrium of stage game $\Gamma(t)$ is in dominant strategies and is induced by the profile*

$$a^*(t) = \left(a_i^*, i \in I | a_i^* = \mathbb{1}_{\{\pi_i(1,0,t) \geq \pi_i(0,0,t)\}}\right).$$

Lest the special case described in Corollary 1, strategic complementarity engenders multiple Nash equilibria in the stage game.[5] This means that one can construct a great many subgame-perfect equilibria of the infinitely repeated game. To obtain sharper predictions on equilibrium play, I focus on the subgame-perfect equilibrium that maximizes total payoffs across countries. Formally, this equilibrium is selected when applying the refinement of strongly renegotiation-proof equilibrium.[6]

Definition 1. *(Farrell and Maskin, 1989) A subgame-perfect equilibrium s is weakly renegotiation-proof (WRP) if there do not exist continuation equilibria s^1, s^2 of s such that s^1 strictly Pareto-dominates s^2. A WRP equilibrium is strongly renegotiation-proof (SRP) if none of its continuation equilibria is strictly Pareto-dominated by another WRP equilibrium.*

Since the stage-game Nash equilibria are ordered, any continuation equilibrium containing dominated Nash equilibria gives countries an incentive to jointly renegotiate towards one that does not contain dominated Nash equilibria. The SRP equilibrium profile is simply given by the unique sequence of undominated Nash equilibria for every stage of the game.

5. Even when $\gamma = 0$, countries may be indifferent about treaty participation in a particular period. I rule out multiple Nash equilibria in this case by assuming that countries choose to cooperate whenever they are indifferent between both actions. This assumption is not very restrictive in that it rules out equilibria that differ from the chosen equilibrium in at most N periods. This is because Assumption 3 implies that potential ties in payoffs cannot last for longer than one period. Moreover, in the empirical framework below it is assumed that payoff functions are subject to a continuously distributed random disturbance (assumption 4 below), so that ties occur with probability zero.

6. This is the definition of renegotiation-proofness commonly used in the literature on the stability of international environmental agreements (*e.g.* Barrett, 1994; Finus and Rundshagen, 1998; Asheim and Holtsmark, 2009). Pearce (1990) reviews alternative definitions of renegotiation-proofness. In Section A of the Online Appendix, I analyse the issue of multiple equilibrium with strategic complementarities in more detail and show that renegotiation-proofness induces a coalition-proof Nash equilibrium at every stage of the game.

A straightforward algorithm to compute this profile starts in the first period in which full participation in the treaty is a Nash equilibrium. Working backwards in time, a country is dropped from the set of signatories if it has an incentive to defect given the set of other signatories. Since the Nash equilibrium monotonically increases over time, this procedure is guaranteed to find the largest treaty (the largest Nash equilibrium) in every period. Theorem 2 below states that this profile indeed constitutes the unique SRP equilibrium of the repeated game.

To obtain a formal representation of this algorithm, consider a Nash equilibrium a^m in period $t_{(m)}$. Suppose that a^m induces exactly m countries to cooperate. Denote by $K_m \subset I$ the set of m cooperators and use the definition of Nash equilibrium to write

$$\Delta\pi_i\left(a^m_{-i}, t_{(m)}\right) \geq 0 \qquad \forall i \in K_m \tag{2.4}$$

$$\Delta\pi_j\left(a^m_{-j}, t_{(m)}\right) < 0 \qquad \forall j \in I \backslash K_m. \tag{2.5}$$

Further, let $t_{(m)} \in T$ denote the earliest period in which a^m is a Nash equilibrium of the stage game. Then there is a country $j_{(m)} \in K_m$ that has the least incentive to cooperate

$$\Delta\pi_{(m)}\left(a^m_{-j_{(m)}}, t_{(m)}\right) \leq \Delta\pi_i\left(a^m_{-i}, t_{(m)}\right) \qquad \forall i \in K_m \tag{2.6}$$

and that prefers not to cooperate in the period just before $t_{(m)}$

$$\Delta\pi_{(m)}\left(a^m_{-j_{(m)}}, t_{(m)} - 1\right) < 0. \tag{2.7}$$

Starting with $m = N$ and working backwards to $m = 1$, the algorithm uses equations (2.4–2.7) recursively to define N sets $K_{(m-1)} \equiv K_m \backslash \{j_{(m)}\}$, along with the action profiles inducing them $a^m \equiv \left(a_j, j \in I | a_j = \mathbb{1}\{j \in K_m\}\right)$, as well as N periods $t_{(m)}$. These elements are needed in the following theorem to characterize the SRP equilibrium.

Theorem 2. *If $\gamma \geq 0$, the game G has a unique strongly renegotiation-proof equilibrium $s^* = \{s^*(t)\}_{t=0}^{\infty}$ where*

$$s^*(t) = \left(s^*_{(i)}(t), i \in I | s_{(i)}(t) = \mathbb{1}_{\{t \geq t^*_{(i)}\}}\right).$$

Country $j_{(i)}$ joins the treaty in period $t^*_{(i)}$ and never withdraws from it at any time after that. The sequence $\{t^*_{(i)}\}_{i=1}^{N}$ is given by $t^*_{(N)} = t_{(N)}$ and

$$t^*_{(N-i)} = \min\left[t^*_{(N-i+1)}, t_{(N-i)}\right] \qquad i = 1, \ldots, N-1, \tag{2.8}$$

where $t_{(m)}$ is defined in equations (2.4–2.7) for all $m = 1, \ldots, N$.

Characterizing the equilibrium is particularly easy when $\gamma = 0$ because each country always has a dominant action and this is weakly increasing in t, so that the country never wants to withdraw.

Corollary 2. *If $\gamma = 0$, the unique SRP equilibrium is given by the strategy profile $s^* = \{s^*(t)\}_{t \in T}$ where*

$$s^*(t) = \{s^*_i(t), i \in I | s^*_i(t) = \mathbb{1}_{\{t \geq t^0_i\}}\}$$

and t_i^0 is given by

$$t_i^0 \equiv \{t \in T \mid \Delta\pi_i(0,t-1) < 0 \leq \Delta\pi_i(0,t)\}. \tag{2.9}$$

For the case of $\gamma < 0$, the game is no longer supermodular and hence not much can be said about the N-player game without imposing more structure. By assuming symmetric spillovers, I obtain a framework that—while still complex—yields sharp predictions on equilibrium play for the subsequent empirical analysis.

Theorem 3. *If $\gamma < 0$ and $w_{ij} = c \, \forall i,j \in I$ s.th. $i \neq j$ the stage game $\Gamma(t)$ has a Nash equilibrium in pure strategies for all $t \in T$. The Nash equilibrium is unique, up to the identity of treaty members, and is weakly increasing in t.*

Multiplicity of equilibrium only arises if two countries have very similar payoffs to cooperation so that either one—but not both—can be treaty members in a Nash equilibrium. These equilibria are not necessarily Pareto-ranked, so that either one can be part of the SRP equilibrium of the supergame.

Theorem 4. *If $\gamma < 0$, $w_{ij} = c \, \forall i,j \in I$ s.th. $i \neq j$, and countries get to decide on ratification in descending order of their net benefit from treaty participation, ϕ_i, then the strategy profile $s^* = \{s^*(t)\}_{t=0}^{\infty}$ induces a unique strongly renegotiation-proof equilibrium of game G, where*

$$s^*(t) = \left(s^*_{(i)}(t), i \in I \mid s_{(i)}(t) = \mathbb{1}_{\{t \geq t_{(i)}\}}\right).$$

Country $j_{(i)}$ joins the treaty in period $t_{(i)}$ and never withdraws from it at any time after that, where $t_{(i)}$ is defined in equations (2.4–2.7) for all $i = 1, \ldots, N$.

To establish uniqueness of the SRP equilibrium, I resolve multiplicity of equilibrium in the stage game by imposing sequential moves in the stage game, paired with a particular order of moves. This trick has been widely applied in the literature on entry games. It is no longer needed below when I consider asymmetric countries and let the period length go to zero.[7]

The intuition underlying the SRP equilibrium is as follows. Since payoffs are monotonically increasing with time, full cooperation is a dominant-strategy equilibrium of the stage game from some period $\hat{t}$ onwards (specifically, $\hat{t} = t_N^0$ for $\gamma \geq 0$ and $\hat{t} = t_{(N)}$ for $\gamma < 0$). Trivially, this continuation equilibrium is WRP because no other profile is played along the equilibrium path. It is also the unique SRP continuation equilibrium because full cooperation strictly dominates every other stage-game equilibrium. Since no punishment of prior deviations is possible beyond period $\hat{t}$, a Nash equilibrium of the stage game must be played in all earlier periods. For $\gamma = 0$, this gives rise to a unique sequence of stage-game Nash equilibria in dominant strategies. For $\gamma > 0$, the stage game may have multiple Nash equilibria but they are Pareto-ranked. Again, only the largest Nash equilibria in each stage game can be part of the SRP equilibrium. All that remains to be shown is that the algorithm correctly identifies the largest Nash equilibrium in every stage of the game. This step is proven by induction exploiting the monotonicity of Nash equilibrium with respect to time. For $\gamma < 0$ and symmetric spillovers, there Nash equilibrium is either unique or all Nash equilibria are equally sized, in which case the algorithm picks the one induced by the assumed order of moves.

7. See the remark following the proof of theorem 4 in the Appendix.

3. EMPIRICAL IMPLEMENTATION

3.1. *Implications for empirical analysis*

A comparison of the equilibrium outcomes characterized in the previous section reveals the effects of strategic interaction on the timing of treaty participation. Strategic substitutability does not affect the first country to ratify the treaty, but it delays subsequent ratification times. In contrast, strategic complementarity accelerates participation by reinforcing the incentive to cooperate. Moreover, if this effect is sufficiently strong, treaty ratification by one country triggers ratification by others, leading to clustered ratification decisions, as depicted in Figure 1 above.

In order to exploit this prediction in an empirical application to learn about the sign and strength of strategic interaction effects, one must rule out reasons for clustering that may occur even in the absence of strategic complementarity. For example, as the length of a time period increases from a day or week to a month or year, it is inevitable that ratification times of some countries are clustered. To abstract from this effect, I characterize the equilibrium as time periods become infinitesimally small. A first step into this direction is to write down the continuous-time analogues to the equilibrium ratification times in the discrete-time game. Define $\tilde{t}_i^0 = \{\tilde{t} \in \mathbb{R}_+ | \Delta\pi_i(\mathbf{0}, \tilde{t}) = 0\} = g^{-1}(\phi_i)$ as the first calendar time at which cooperation is a dominant strategy for country i, and let $\tilde{t}^*_{(N)} = \min\{\tilde{t} \in \mathbb{R} | \Delta\pi_i(\mathbf{1}, \tilde{t}) \geq 0 \forall i \in I\} = \max_i \left[g^{-1}\left(\phi_i - \gamma \sum_{j \neq i} w_{ij}\right)\right]$ denote the first calendar time at which full participation is a Nash equilibrium of the stage game. Computation of the remaining calendar times $\tilde{t}^*_{(N-1)}, \ldots, \tilde{t}^*_{(1)}$ in a SRP equilibrium is by straightforward application of the algorithm described in Section 2.3 above.[8]

To establish the link between ratification events in discrete and continuous time, consider the condition

$$\Delta\pi_i(a_{-i}, t) \geq 0 \tag{3.13}$$

where a induces some set of member countries. For example, $a = \mathbf{0}$ pins down t_i^0 and $a = \mathbf{1}$ pins down $t^*_{(N)}$. The only difference between discrete periods $t^* \in T$ and calendar time $\tilde{t}^* \in \mathbb{R}_+$ is that the latter solves this condition with equality (due to the continuity Assumption 3 and by the intermediate value theorem) whereas in the discrete-time game one needs to find the first period on the time grid satisfying the weak inequality. On a time grid with period length one the relationship between the two is thus given by $t^* = \lceil \tilde{t}^* \rceil$.

Consider now the sequence of discrete-time games $\{^kG\}$, where ^kG is played over a grid with period length 2^{-k}, and $k \in \mathbb{N}_0$. A node on this grid, ^kt, has the property $^kt = t2^k$. That is, an increase in k by 1 doubles the number of decision nodes and cuts the period length in half. As k goes to infinity, the grid length 2^{-k} goes to zero. The SRP equilibrium in the limit game is characterized by the following theorem.

8. Some adjustments to the notation are needed. Denote by subscript (m) the player $j_{(m)} \in I$ who is just indifferent between ratifying or not in the stage game played at time $\tilde{t}_{(m)}$ if all countries in K_m participate. Compute this time as

$$\tilde{t}_{(m)} = \{\tilde{t} \in \mathbb{R} | \Delta\pi_{(m)}(a^m_{-j_{(m)}}, \tilde{t}_{(m)}) = 0 \wedge \Delta\pi_i(a^m_{-i}, \tilde{t}_{(m)}) \geq 0 \forall i \in K_m \wedge \Delta\pi_j(a^m_{-j}, \tilde{t}_{(m)}) < 0 \forall j \notin K_m\}. \tag{3.10}$$

Define

$$K_{m-1} \equiv K_m \backslash j_{(m)}. \tag{3.11}$$

Let $K_N = I$ and compute $\tilde{t}_{(m)}, j_{(m)}, K_m$ by recursive application of (3.10) and (3.11) for $m = N, N-1, \ldots, 1$. The equilibrium ratification times are given by $\tilde{t}^*_{(N)} = \tilde{t}_{(N)}$ and

$$\tilde{t}^*_{(N-i)} = \min\left[\tilde{t}_{(N-i+1)}, \tilde{t}_{(N-i)}\right] \qquad i = 1, \ldots, N-1. \tag{3.12}$$

Theorem 5. *As time periods become infinitesimally small, the SRP equilibrium of the limit game $^{\infty}G$ exists and equilibrium provision times converge to their continuous-time analogues*

$$\lim_{k\to\infty} {}^{k}t_i^* \cdot 2^{-k} = \tilde{t}_i^* \qquad \forall i \in I.$$

With this result in hand, it is straightforward to establish that clustering of ratification times is a probability-zero event if countries have different private net benefits of treaty ratification:

Assumption 4. *For all $i \in I$, $\phi_i \in \mathbb{R}$ is a continuous, i.i.d. random variable with distribution function $F(\phi)$.*

Theorem 6. *Under Assumptions 1—4, if time periods become infinitesimally small then clustering of ratification decisions among asymmetric countries occurs only in the presence of strategic complementarity.*

The theorem suggests an intuitive strategy to empirically identify strategic complementarity in the timing of treaty participation: given a data set in which countries start out as non signatories to a treaty and participate only at a later stage, the timing of the ratification decision traces out the distribution of the relative benefits to cooperation. Furthermore, if ratification times are recorded on a sufficiently fine time grid, clustering must be attributed to strategic complementarity unless countries are identical.

3.2. *Parameter identification*

In line with Assumptions 3 and 4, I assume an exponential trend $g(t) = e^{\lambda t}$ and an exponential index of net cost ϕ_i=exp($x_i'\beta$+ϵ_i) where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon)$ is a vector of *i.i.d.* disturbances that are observed by all players but not by the econometrician. The vector x_i contains country characteristics that shift the private net benefit of ratification. Assumption 1 is satisfied for $r_i > \lambda$. The relative payoff to ratification takes the form

$$\Delta\pi_i(a_{-i}, t) = -\exp(x_i'\beta + \epsilon_i) + e^{\lambda t} + \gamma \textstyle\sum_{j \neq i} w_{ij} a_j(t). \tag{3.14}$$

Applying the results from Section 2, I compute the equilibrium ratification times $\{t_{(i)}^*\}_{i \in I}$ using the following algorithm:

Algorithm. *Let K be a set of signatories including country i. Denote by $\tilde{t}(K) \in \mathbb{R}^N$ the vector of calendar times with elements*

$$\tilde{t}_i(K) = \frac{1}{\lambda} \log\left[\phi_i - \gamma\left(\sum_{j \in K\setminus\{i\}} w_{ij}\right)\right] \quad , i \in K. \tag{3.15}$$

These are potential equilibrium ratification times of countries in K. Compute the identity of the last country to ratify the agreement as

$$j_{(N)} \equiv \arg\max_{k \in I} \{\tilde{t}_k(I)\}$$

and $j_{(N)}$s ratification time as $t_{(N)}^* = \lceil \tilde{t}_{(N)}(I) \rceil$.

Recursively, for $m = N-1, N-2, \ldots, 1$*, define the set* $K_m \equiv K_{m+1} \backslash j_{(m+1)}$ *and compute the identity of the mth country to join the agreement as*

$$j_{(m)} \equiv \arg\max_{k \in K_m} \{\tilde{t}_k(K_m)\}$$

with ratification time $t^*_{(m)} = \min\left[t^*_{(m+1)}, \lceil \tilde{t}_{(m)}(K_m) \rceil\right]$.

The vector of model parameters to be estimated is given by (γ, β, λ). For fixed $\sigma_\epsilon = 1$, the parameter vector β and λ are parametrically identified from the joint distribution of country characteristics and ratification times unless all countries ratify on the same day. [9] For $\gamma = 0$, a country's optimal ratification time is given by $t_i^0 = \lceil \frac{1}{\lambda} \ln(\phi_i) \rceil$, *i.e.* the smallest integer to satisfy $-\phi_i + e^{\lambda t} \geq 0$. This sweeps out the distribution of ϕ_i. For $\gamma \neq 0$, the timing of ratification is also determined by the magnitude of the net spillovers γw_{ij} between any two countries i and j. There are three sources of identification of the spillover parameter γ. First, for a given parametric specification, the time elapsed between subsequent ratification events is informative about the sign and magnitude of γ. Secondly, the presence of ratification clusters in the data identifies strategic complementarity regardless of the specification. Both these sources of identification are available when testing for the sign of γ under the assumption that spillovers are homogenous. Thirdly, under the assumption that $\gamma \geq 0$, data on spillover intensities w_{ij} between country pairs (i, j) can be exploited to test whether strategic complementarity arises through a particular bilateral channel, such as trade relationships, reputation effects, or equity concerns. In sum, this suggests an analysis in two steps. In the first step, γ is estimated under the assumption of homogenous spillovers $w_{ij} = \frac{1}{N-1}$, $i \neq j$ and $w_{ii} = 0$. If this yields a positive estimate $\hat{\gamma}$, a second estimation substitutes heterogeneous spillovers for w_{ij} and proceeds under the constraint that $\gamma \geq 0$.

This identification strategy rests on two key assumptions. First, Assumption 4 guarantees that players are not identical with probability one, as the random disturbance shifts the payoff function in a strictly monotonic fashion. This assumption is frequently made in applied research to prevent "over fitting" of the model. Secondly, I assume a functional form of $\Delta\pi$ with a smooth and additive time effect $g(t)$. If, contrary to this assumption, the relative payoff to cooperation were subject to a common, positive shock, then two or more countries that are close to the threshold will ratify immediately after the shock hits, even as the grid length goes to zero. This would lead to an overestimation of the magnitude of strategic complementarity. Conversely, a sudden drop in relative benefits to cooperation would result in an underestimation of strategic complementarity or overestimation of strategic substitutability. Assumption 3 rules out common shocks. Whether or not this is reasonable depends on the particular application at hand and will be further discussed in context below.

3.3. *Econometric estimation*

Since the expectation of equilibrium ratification times does not have a closed-form solution in the presence of strategic complementarity, I employ the method of simulated moments (MSM) to estimate the parameter vector $\theta = (\beta, \gamma, \lambda)$. The estimation algorithm takes S random draws ϵ^s on the distribution of ϵ and solves for the vector of ratification times, $t(\epsilon^s; \theta_0)$, for a given vector of parameter values θ_0. A consistent estimator of θ is given by

$$\hat{\theta} = \min_\theta g(\theta)' W g(\theta) \tag{3.16}$$

9. The statistical population considered here is comprised of individual countries. Alternatively, when a cross-section of identical treaties is available, treaties can be considered as forming the statistical population.

where

$$g = \frac{1}{N}\sum_{i=1}^{N}\left(\mu(t_i) - \frac{1}{S}\sum_{s=1}^{S}\mu(t_i(\epsilon^s;\theta))\right) \times f(X_i) \tag{3.17}$$

is a vector of sample analogues to the moment conditions μ of observed and simulated ratification times t_i and $t_i(\epsilon^s;\theta)$, respectively (see McFadden, 1989; Pakes and Pollard, 1989; Lee and Ingram, 1991). The term $f(X)$ denotes functions of the instruments X. The estimator $\hat{\theta}$ converges in probability to θ and $\sqrt{N}(\hat{\theta}-\theta)$ converges in distribution to a normally distributed random vector with mean zero and covariance matrix

$$\left(1+\frac{1}{S}\right)\left[E_0\frac{\partial g'}{\partial \theta}W^{-1}\frac{\partial g}{\partial \theta'}\right]^{-1} E_0\frac{\partial g'}{\partial \theta}W^{-1}E_0 gg' W^{-1}\frac{\partial g}{\partial \theta'}\left[E_0\frac{\partial g'}{\partial \theta}W^{-1}\frac{\partial g}{\partial \theta'}\right]^{-1}. \tag{3.18}$$

Using the optimal weighting matrix $W^* = E_0 gg'$ the asymptotic covariance matrix of $\sqrt{N}(\hat{\theta}-\theta)$ can be reduced to

$$\left(1+\frac{1}{S}\right)\left[E_0\frac{\partial g'}{\partial \theta}W^{*-1}\frac{\partial g}{\partial \theta'}\right]^{-1}. \tag{3.19}$$

The estimation algorithm computes the simulated moments and evaluates the criterion function (3.16) with W taken to be the identity matrix. Minimization of equation (3.16) in this fashion yields an initial consistent estimate of θ that is used to compute a consistent estimate of W^*. Standard errors are obtained via non-parametric bootstrap with re-sampling and 150 random samples.

The moments are chosen so as to capture relevant variation in the data that helps to identify the parameters of the model (Section D.2 of the Online Appendix provides more details). For the minimization of the MSM criterion function (3.16) I sequentially employ two algorithms. First, I use Goffe *et al.*'s (1994) implementation of the simulated annealing algorithm to avoid that the algorithm gets trapped in a local minimum. After the simulated annealing algorithm has closed in on a minimum, the result is turned over as the starting value to a Nelder–Mead simplex algorithm which provides the final vector of parameter estimates.

4. INTERNATIONAL ENVIRONMENTAL AGREEMENTS: THE MONTREAL PROTOCOL

4.1. *Background*

The objective of the Montreal Protocol is to phase out emissions of CFCs and other man made pollutants that deplete ozone molecules in the stratosphere.[10] Since their invention in the 1920s, CFCs were widely used as freezing agents in air conditioning, as aerosol propellants in spray cans, as plastic-foam-blowing agents and as solvents in the manufacturing of microchips and telecommunications parts. The soaring demand for these products fueled a very rapid growth in worldwide CFC production that lasted until the 1970s, when atmospheric chemists predicted that unaltered CFC emissions would substantially reduce ozone concentrations in the stratosphere. This was a reason for concern because stratospheric ozone acts as a shield against harmful UV B radiation from space which increases the prevalence of skin cancer and eye cataracts in humans and causes damage to crops and ecosystems. By one estimate, fatal cases of skin cancer in the U.S. population would increase by more than 3.1 million between 1985 and 2075 in the absence

10. Section B of the Online Appendix provides further details on ozone depletion and the Montreal Protocol.

of further regulation of CFC emissions (EPA, 1987). The empirical relevance of stratospheric ozone depletion became very salient in 1985, when new data showed what became known as the "ozone hole"—a seasonal decline in ozone concentrations of about 50% compared to levels measured in the 1960s. Subsequently, an international expert panel concluded that worldwide ozone losses were wholly or in part due to CFCs.

Since CFCs mix uniformly in the stratosphere, policy-makers recognized the public-good nature of CFC abatement and that its efficient provision would require international cooperation. An international treaty, the Montreal Protocol on Substances that Deplete the Ozone Layer, was negotiated and opened for signature on 16 September 1987. This treaty stipulated a 50% cutback in consumption of five CFCs from 1986 levels by 1998, with a stabilization by 1989 and a 20% reduction by 1993. It also enacted a freeze of three halons at 1986 levels. Under article 5 of the Protocol, developing countries with a per-capita consumption of less than 0.3 kg were granted a grace period of 10 years to meet these obligations. To prevent leakage, the treaty prohibited bulk imports of controlled substances from non-Member States and bulk exports from developing countries. Trade between developed member countries was left unrestricted, but imports were counted towards a country's net consumption.[11]

The Montreal Protocol went into force on 1 January 1989, by which time it had been ratified by most major developed countries. To encourage accession by large developing countries such as India and China, the second meeting of parties held in London in June 1990 established a multilateral fund to pay for the incremental costs of compliance incurred by these countries. Contributions to the fund were the responsibility of developed country parties according to the UN scale of assessment. Subsequent meetings brought forward the time table for phasing out ozone-depleting substances, augmented the list of controlled substances and adopted measures to curb illegal trade. The global phase-out of CFCs has been proceeding swiftly within the parameters set by the Montreal Protocol, making the treaty one of the most successful international environmental agreements so far.

4.2. *Assumptions*

The empirical framework developed above nests the competing hypotheses of complementarity, substitutability and dominance of ratification decisions. This is well suited to capture the multifaceted and fundamentally asymmetric incentives to participate in the Montreal Protocol. The most salient aspect of the treaty is the public good of protecting the ozone layer. This is reflected in Assumption 2. The payoff specification flexibly admits both constant marginal external benefits of abatement ($\gamma_0 = \gamma_1$) or decreasing marginal external benefits ($\gamma_1 < \gamma_0$). The model also admits the possibility of strategic complementarity ($\gamma_0 < \gamma_1$) which may arise as a consequence of the trade ban, issue linkage, reputation effects or social preferences, as will be discussed in Section 4.5 below.

Since ozone depleting substances accumulate in the stratosphere, the underlying economic problem is one of controlling a stock pollutant with multiple agents. While it would be desirable to fully account for the dynamics of stock pollution, to my best knowledge, the theoretical literature on multilateral agreements over stock pollution has not yet developed models that are amenable to econometric estimation.[12]

11. Art. 1 (6) of the Montreal Protocol defines a country's net consumption as production plus imports minus exports of controlled substances.Targets had to be met as a weighted average across all controlled substances, with weights corresponding to their "ozone-depleting potential".

12. This literature uses differential game theory to analyse multilateral agreements for controlling stock pollutants among N countries. In order to be able to calculate a numerical solution to such a model, Rubio and Ulph (2007) assume

Therefore, I resort to simpler dynamics, governed by the function $g(t)=\exp(\lambda t)$, which are consistent with an exogenous process—such as GDP growth—driving participation. Using the standard model of the voluntary provision of a public good, Murdoch and Sandler (1997) have argued that income growth was the principal driver of CFC abatement before the Montreal Protocol went into force. This income effect predicts that, as countries grow richer over time, more of them are willing to assume the cost of complying with the abatement targets mandated under the Montreal Protocol. Modelling GDP growth as an exponential trend is standard. Additivity and the uniform growth rate λ are simplifying assumptions made to ensure tractability of the model.[13]

As was explained in Section 3.1 above, identification of γ hinges on the assumption that there are no common shocks to $\Delta\pi$. For example, a sudden cost drop that occurs only in country i gets absorbed into the estimate of ϕ_i while a smooth decline in the cost of CFC substitutes gets absorbed into the estimate of λ.[14] Both leave the estimate of γ unaffected. In contrast, an unanticipated drop in the cost of CFC abatement that pushes several countries over the threshold causes clustering of ratification decisions and hence increases the estimate of γ.

In principle, the assumption of a smooth trend in the net benefits to ratification is not testable. In view of an analysis of CFC abatement cost in the U.S.—the world leader in CFC production at the time—by Hammitt (2000) the possibility of unanticipated drops appears rather unlikely. Hammitt documents that inflation-adjusted wholesale prices for CFCs increased monotonically over time (his data cover the years from 1986 until 1994). Had prices fallen during this time period, unanticipated drops in marginal abatement cost would seem to be more of a concern. In fact, Hammitt shows that contemporary estimates of marginal abatement cost published by EPA in December 1987 and August 1988 *underestimated* the realized cost. This suggests that, if anything, countries were expecting actual abatement cost to be too low and would have had an incentive to postpone ratification once they found out. Since this would lead to downward bias in the estimated γ, the results obtained for $\hat{\gamma} > 0$ below can be interpreted as conservative estimates of strategic complementarity.

4.3. *Data*

The data are taken from various sources and summarized in Table D.1 in the Online Appendix. The endogenous variable is the day on which a country ratified the Montreal Protocol, which marks the legal act of acceding to the agreement. Ratification dates are available at UNEP's ozone web site.[15] The mean time to ratification in the sample of 140 countries is 5 years, 2 months, and 10 days. More than 20 years elapsed between the first and last ratification (see also Figure D.1 of the Online Appendix). This suggests that the net benefits of joining were highly dispersed and provides ample variation for the estimation of the empirical model. Moreover, while ratification by most countries is spaced out over time, some ratified the treaty in clusters, *e.g.* in the same month or week, or even on the same day.

The choice of covariates in the vector x is subject to a tradeoff. Ideally, one would like to include all variables that could possibly shift a country's net benefits. In practice, however, the

that all countries are identical. A recent strand of theoretical research blends contract theory and dynamic games to analyse the optimal design of agreements over stock pollutants (Harstad, 2012, 2015; Battaglini and Harstad, 2015).

13. An alternative justification for the monotonicity Assumption (3) is that, in a world with a monopoly country producer, that country would eventually find it worthwhile to curb its use of CFCs and switch to a "backstop technology." I am grateful to an anonymous referee for suggesting this argument.

14. Recall that the empirical model does not discriminate between smooth reductions in costs and smooth growth in the benefits of abatement over time.

15. See http://ozone.unep.org (last accessed on January 15, 2016).

need to estimate a highly nonlinear model with a relatively small data set dictates a parsimonious specification that conserves on degrees of freedom and minimizes multi-collinearity.[16] The most severe known consequence of the depletion of the ozone layer is an increased risk of skin cancer. Per capita income and population (both in logs) are thus expected to shift outward the net benefit of abatement in per capita and in absolute terms, respectively. Since stratospheric ozone is thinning more quickly in higher latitudes, the benefit of CFC abatement should be higher there. Whether or not a country hosts a CFC producing firm is expected to shift the cost of CFC abatement. A dummy for article 5 countries picks up structural differences between developing and developed countries, as well as the effect of the grace period granted to the latter. Furthermore, I include a dummy variable for article 5 countries that joined after the London Amendments were negotiated in June 1990. This controls for unanticipated changes in the incentive to participate which resulted from the offer of financial aid to developing countries that acceded to the treaty.

All covariates are evaluated at their 1986 levels—the year before the treaty was opened for signature—in order to preclude simultaneity issues. Country characteristics such as GDP in 1995 U.S.$ and population size in millions were taken from the World Development Indicators.[17] The producer dummy equals 1 for all countries that report positive production of any of the five CFCs regulated under the Montreal Protocol, based on production data taken from UNEP (2015). The list of countries with article 5 status was obtained from the UNEP web site. The latitude variable is taken from the CIA World Fact Book and refers to the location of the country's capital.[18]

4.4. *Global interaction effects*

Table 1 reports estimates of the parameter vector (γ, β, λ) when symmetric weights $w_{ij} = (N-1)^{-1}$ are imposed. The parameter γ can thus be interpreted as a global strategic interaction effect and can take on either positive or negative values. The first three columns of the table report different the estimates obtained for different sets of covariates. The estimated γ is positive and statistically significant for all specifications, indicating that ratification decisions are strategic complements. Columns 1-3 show that the point estimate for γ diminishes slightly and becomes more precise as further covariates are added to the model.

The estimated coefficients on the variables shifting the *net cost* of accession are robust across specifications and in line with intuition. For example, the negative coefficient on per capita income suggests that richer countries place a greater value on environmental quality. Similarly, the negative coefficient on population indicates that populous countries benefit more from global abatement efforts in absolute terms than less populous countries. The coefficient on the dummy variable for article 5 countries is negative and suggests that the 10-year grace period for abatement lowered the cost of accession. In contrast, the dummy variable ART5xLONDON must, by construction, enter with a positive coefficient in the cost term. The coefficient on CFC production is positive and reflects the fact that producer countries were facing higher compliance costs, ceteris paribus. The point estimate for latitude is statistically insignificant.

16. For example, the final specification includes a dummy variable for CFC producing countries and not the country's actual CFC consumption because the latter is highly collinear with income, population and article 5 status. Table D.2 in the Online Appendix lists all included covariates along with the rationale behind their inclusion and the expected signs of their coefficients.

17. Available online at http://data.worldbank.org/data-catalog/world-development-indicators (last accessed on January 18, 2016).

18. Available online at https://www.cia.gov/library/publications/the-world-factbook/ / (last accessed on January 18, 2016). The variable is reported in hundreds of degrees for computational reasons.

Both the estimation results for γ and the observed clustered ratification events in the data lend empirical support to the fundamental decision-making process proposed in this article.[19] A crucial aspect for the plausibility of this mechanism is that a country correctly anticipates ratification by another, so that both can plan ahead and try to coordinate the ratification date in mutually beneficial ways.[20]

In the real world, do countries—especially democratic ones—have the ability to actually coordinate on the same ratification date, as posited by the model? Benedick (1998, p. 116f) describes a compelling case-in-point. Eight out of twelve Member States of the EU (then EC) ratified the Montreal Protocol simultaneously on 16 December 1988. Two other Member States ratified two weeks later, and the other two countries had ratified the Protocol already two months earlier. This coordination was achieved by the European Commission asking its Member States, many months before, to ratify the Montreal Protocol simultaneously with the Commission. Obviously, this coordination device was not perfect, but it worked for the majority of countries, and despite the fact that the process had to be approved by eight national parliaments.

The coefficient estimates in columns 1-3 are based on the assumption that countries coordinate on the Pareto-efficient outcome at every stage of the game—embodied in the equilibrium refinement of strong renegotiation-proofness. In view of the EC example, one might wonder what an overly optimistic assumption about coordination implies for the point estimate of γ. To investigate this, I re-estimate the model based on the subgame-perfect Nash equilibrium with complete coordination failure. In the simple two-player game shown in Figure 1(c), this means that players start to cooperate in the latest possible period t_1^0 rather than in the earliest possible period t_2^1. Imposing coordination failure more than triples the point estimate γ in column 4 of Table 1. The reason is that, as coordination on the early period fails, strategic complementarity must be stronger to rationalize a given speed of ratification in the data. This is why the preferred estimate in column 3 can be viewed as a lower bound on the magnitude of global strategic complementarity.

How strong are the strategic complementarities implied by the estimated γ parameters? The structural model allows me to answer this question by comparing the predicted ratification sequence that arises with strategic interaction to the sequence of ratification that results when the strategic interaction effect is removed. The latter is obtained in a counterfactual experiment that sets $\gamma = 0$ and simulates the ratification times based on the parameter estimates for the remaining model parameters. Panel B of Table 1 reports the mean (median) difference between the ratification times obtained in the counterfactual experiment without strategic complementarity and those predicted by the fitted model with $\hat{\gamma} > 0$. In the preferred specification reported in column 3, strategic complementarity accounts for a reduction by 208 days (190 days) in the mean (median) ratification time. Relative to the non-strategic counterfactual, strategic complementarity accelerated the mean (median) ratification time by 12% (9%). This effect is economically significant. Under the alternative equilibrium selection rule where coordination fails, the reduction in the mean (median) time to ratification is even larger and amounts to 22% (17%).

4.5. *The sources of strategic complementarity*

Having established that global interaction effects in the ratification of the Montreal Protocol are strategic complements rather than substitutes, I now turn to studying the sources of this effect. The

19. Sections D.2 and D.3 of the Online Appendix analyse in detail the goodness of fit and the robustness of these results to changing the specification, the functional form of the time trend and the sample size.

20. As illustrated in Figure 1(d), when strategic complementarity is strong there is usually a window of opportunity for coordinating ratification times.

TABLE 1
Parameter estimates: Global interaction effect

	(1)	(2)	(3)	(4)
A. Parameter estimates				
γ	4.665	3.874	3.398	11.665
	(1.581)	(1.041)	(1.072)	(2.967)
Constant	3.813	4.069	3.828	3.343
	(0.430)	(0.369)	(0.386)	(0.163)
Log income p.c.	−0.285	−0.319	−0.302	−0.165
	(0.055)	(0.048)	(0.049)	(0.033)
Log population	−0.179	−0.188	−0.184	−0.154
	(0.042)	(0.043)	(0.042)	(0.039)
CFC producer (D)		0.042	0.015	0.019
		(0.021)	(0.005)	(0.005)
Latitude			0.047	−0.560
			(0.025)	(0.168)
Art. 5 country (D)	−0.259	−0.320	−0.253	−0.215
	(0.093)	(0.106)	(0.088)	(0.052)
Art. 5 × post-London (D)	1.327	1.333	1.318	1.406
	(0.160)	(0.159)	(0.161)	(0.147)
$\lambda \cdot 10^4$	10.123	10.094	9.934	10.881
	(1.115)	(1.118)	(1.080)	(1.118)
B. Acceleration of ratification times w.r.t. non-strategic counterfactual				
Δ mean (days)	263	226	208	411
(%)	15	13	12	22
Δ median (days)	233	205	190	391
(%)	11	10	9	17
Coordination	yes	yes	yes	no

Notes: $N = 140$ countries. Spillovers are symmetric $w_{ij} = \frac{1}{N-1} \forall i \neq j$ and $w_{ii} = 0 \forall i$. Bootstrapped standard errors are in parentheses.

economics and international relations literature on international treaties have discussed specific government interaction effects that engender strategic complementarity. It has been suggested that governments (or their constituents) have preferences for fair and equitable sharing of the abatement burden, or that social norms engender a reputation loss for non-signatories. It has also been pointed out that linkage of environmental cooperation to other international policy issues, or to international trade, can deter free-riding. In what follows, I briefly review how these effects induce strategic complementarity in the context of the Montreal Protocol (a more detailed review of this is relegated to Section C of the Online Appendix) and use the empirical framework developed above to test whether they matter empirically. This is accomplished by using actual data on bilateral spillovers to calibrate w_{ij} in equation (2.2) before estimating the parameter $\gamma \geq 0$ associated with a particular mechanism of interaction between governments.

4.5.1. Economic dependency and the trade ban. Previous research has conjectured that a country's ratification decision may be sensitive to the behaviour of countries upon which it is economically dependent. Using data on the early phase of the Montreal Protocol, Beron *et al.* (2003) test whether a country i is more likely to follow suit to another country j's ratification if a large share of i's exports go to j. I adopt their "power matrix" by computing spillover weights as

$$w_{ij} = \frac{\text{exports from } i \text{ to } j}{\text{total exports from country } i}. \tag{4.20}$$

I compute these weights as the average export shares between 1980 and 1986, using bilateral flows of total commodity exports from the NBER-United Nations Trade Data (Feenstra *et al.*, 2005).[21]

Another reason for including trade-based measures of interdependencies is the treaty's ban of trade in controlled substances between parties and non-parties. Barrett (1997b) argues that this ban transformed trade in controlled substances into a club good, the benefits of which were exclusive to Member States. Once the club reaches a critical size, trade with Member States gives higher benefits than trade with non-Member States (cf. Section C.2.4 of the Online Appendix). To gauge the effect of trade restrictions on ratification, one can thus consider the bilateral export shares of controlled substances before the Montreal Protocol went into force. While the UN Ozone Secretariat has a mandate for collecting these data, it is not allowed to publish them. Therefore, I approximate trade in controlled substances using bilateral export flows for a set of four-digit industries that either produce controlled substances or that rely heavily on those substances as inputs. In selecting these industries, I follow UNEP's guidelines for Member States regarding the use of customs data to comply with reporting requirements UNEP (1999) (cf. Figure D.2 in the Online Appendix).

4.5.2. Issue linkage and reputation effects.

Another mechanism for creating complementarity is by linking different policy issues. Diplomats may choose to negotiate different topics jointly in order to achieve more stable outcomes (Tollison and Willett, 1979; Sebenius, 1983). Folmer *et al.* (1993) and Cesar and de Zeeuw (1996) provide specific examples of how such interconnections can help to stabilize an international environmental agreement. Essentially, the greater the number of policy issues in which two countries are involved the better the prospects for linking those issues in a mutually beneficial way. In order to measure this effect, I calibrate w_{ij} to the degree to which country pair i,j is also involved in one of R pre-existing international agreements,

$$w_{ij} = \frac{1}{R(N-1)} \sum_{r=1}^{R} \mathbb{1}\{i \text{ and } j \text{ signed treaty } r\}. \tag{4.21}$$

These weights are computed using data collected by Hathaway (2002) on membership in eleven eminent international agreements. I focus on pre-existing treaties and use membership status as of 1986 in order to avoid potential simultaneity between the decision to join the Montreal Protocol and to ratify other treaties.

Closely related to issue linkage is the notion that states ratify international agreements out of a desire to conform with other countries. For instance, Hoel and Schneider (1997, p. 155) argue that "a government may feel uncomfortable if it breaks the social norm of sticking to an agreement of reduced emissions, even if in strict economic terms it may benefit from being a free rider". In the present model, their argument implies that country i's reputation benefit from ratification is greater the more of its "peers" are among the signatories.[22] Hence, the treaty weights (4.21) can also be interpreted as a plausible (though not the only conceivable) proxy for the reputation cost of breaking a social norm.[23]

21. Section D.1 of the Online appendix provides more details on the computation of these and other spillover weights.

22. In a binary choice framework, a cost exclusive to non-members is equivalent to the benefit $\gamma \sum w_{ij}$ exclusive to Member States.

23. Fundamentally, it remains challenging to distinguish empirically between issue linkage and social norms because it might well be the implicit threat of retaliation in various policy domains which enforces social norms at the intergovernmental level.

4.5.3. Fairness and equity. Preferences for fairness and equity have always been pre-eminent in the public debate on international environmental agreements and have shaped many such treaties in one way or another. For example, most agreements on transboundary pollution stipulate uniform percentage reductions in emissions because they appear to be more equitable than differential abatement targets. Fairness has received considerable attention in the recent economics literature on social preferences (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), and has been given some consideration in the literature on environmental agreements as well (Hoel, 1992; Lange and Vogt, 2003). Here I allow for the possibility that the decision to ratify the Montreal Protocol is at least partially driven by concerns about fairness and equity. In particular, I conjecture that a treaty is perceived as more equitable the more large polluters have joined it, where the size is measured as the share in global CFC emissions

$$w_{ij} = w_j = \frac{(\text{CFC emissions})_j}{\sum_{k=1}^{N} (\text{CFC emissions})_k}. \tag{4.22}$$

I use data on CFC consumption from UNEP (2015) to compute these weights. Consumption is calculated as production plus imports minus exports of controlled substances and is reported by the Member States to the treaty secretariat which monitors compliance. Consumption is measured in metric tons and comprises all five CFCs that were regulated under the Montreal Protocol, weighted by their ozone-depleting potential.

A positive coefficient γ means that accession of large emitters to the Montreal Protocol accelerates ratification by other countries. In order to distinguish the effect of social preferences on ratification from other size effects, I control for GDP using weights defined analogously

$$w_{ij} = w_j = \frac{(\text{GDP})_j}{\sum_{k=1}^{N} (\text{GDP})_k}. \tag{4.23}$$

Since size is a potential confounder of all of the strategic effects discussed here, I include the GDP weights in all specifications.

4.5.4. Results. The sample for which all the weights can be computed contains 103 countries and is summarized in panel B of Table D.1 in the Online Appendix. Compared to the larger sample used above, the average country in this sample is larger, richer, more likely to produce CFCs and tends to ratify earlier.

Table 2 displays the results when the different weights are used in the estimation of the model. Panel A reports the estimates for two interaction coefficients: the first one, γ, corresponds to the particular weight—among the ones discussed above—that is used in the specification of the local interaction effect. The second one, γ_{GDP}, corresponds to the GDP weights in equation (4.23).

The magnitude of the estimated $\hat{\gamma}$ coefficients is not directly comparable across columns because the weighting matrices are calibrated in different ways. For better comparability, panel B reports the acceleration of the mean ratification time that can be attributed to each channel through which strategic complementarity affects ratification times. Using counterfactual analysis, the partial effect of size is computed by subtracting the ratification times of the fitted model with parameters $(\widehat{\gamma}, \widehat{\gamma_{\text{GDP}}})$ from the fitted ratification times in a counterfactual scenario with $(\hat{\gamma}, 0)$. Analogously, I compute the partial effect of the interaction effect measured by γ as the difference between ratification times in the counterfactual model $(0, \widehat{\gamma_{\text{GDP}}})$ and the fitted model. As before, I report these differences both in days and in per cent relative to a non-strategic counterfactual where $\gamma = \gamma_{\text{GDP}} = 0$.

TABLE 2
The sources of strategic complementarity

	(1)	(2)	(3)	(4)	(5)
Weights w_{ij}	Equal	Bilateral exports		Treaty	CFC
		Total	Controlled substances		
A. Coefficient estimates γ					
γ	18.630	3.662	3.964	27.750	6.499
	(7.316)	(1.307)	(2.010)	(10.634)	(3.123)
γ_{GDP}		6.069	8.237	0.001	0.716
		(3.625)	(3.431)	(0.001)	(0.283)
B. Acceleration of mean ratification time w.r.t. non-strategic counterfactual					
Δ in days					
γ	273	185	110	148	392
γ_{GDP}		370	406	0	58
Both		440	432	148	417
Δ in percentage					
γ	19	11	7	9	23
γ_{GDP}		22	24	0	13
Both		26	26	9	25

Notes: $N = 103$ countries. Bootstrapped standard errors are in parentheses.

The first column reports estimates of the benchmark case with equal weights. As before, I find a positive and significant global interaction parameter $\hat{\gamma}$. The implied acceleration in the mean ratification time amounts to 273 days or 19% relative to a non-strategic counterfactual scenario where $\gamma = 0$. The γ parameters obtained using trade weights in column 2 are positive and statistically significant. This means that, on average, countries were more likely to ratify the Montreal Protocol when their principal trading partners were among the members – conditional on the size of the trading partner in the world economy, which is picked up by γ_{GDP}. The effect is also economically significant, as it brings forward the mean ratification time by 11% relative to the non-strategic counterfactual.[24] Following Beron *et al.* (2003), one could interpret this result as the effect of economic dependencies on ratification. Economic dependency measured in this way is necessarily a very broad concept because bilateral trade flows are correlated with a number of country-pair effects, as is shown by the gravity literature.

When focusing on trade in controlled substances and products only, as in column 3, the γ coefficient remains positive and statistically significant, and the implied acceleration of the mean ratification time amounts to 110 weeks (7%) relative to a non-strategic counterfactual. That is, most of the trade-induced strategic complementarity found in column 2 is associated with trade in a narrowly defined segment of controlled substances. This lends empirical support to the hypothesis that banning trade between signatories and non-signatories enhanced participation in the Montreal Protocol (Barrett, 1997b).

Column 4 of Table 2 reports the results obtained when using treaty membership as weights. Again, the estimated γ is positive and statistically significant, implying an acceleration in the mean ratification time of a comparable magnitude as the ban in controlled substances (148 days or 9% of the non-strategic counterfactual). This is in line with the strong positive association

24. Related to this, Neumayer (2002), using a duration model without strategic interaction, finds that export intensive countries ratified the Montreal Protocol earlier.

between ratification of international environmental agreements and membership in international organizations found in reduced-form analysis (Bernauer *et al.*, 2010). The mechanism suggested here is one of local interactions between peer groups of countries that adhere to the same set of social norms and are concerned about their reputation.

Finally, column 5 reports a positive and statistically significant γ-coefficient when simple CFC consumption shares are used as weights. The acceleration of ratification times implied by $\hat{\gamma}$ amounts to 392 days (23%) compared to the non-strategic counterfactual, meaning that ratification of a treaty with targets and timetables for pollution control proceeds more swiftly once the large emitters participate. This result is consistent with social preferences, in that an agreement that includes the big emitters—and, hence, one that appears more equitable to other governments and their political constituencies—enhances the incentive for ratification.

Because CFC consumption across countries is positively correlated with GDP, it is particularly important in this specification to control for strategic interaction that may arise because of the different size of countries and that might confound the estimates of γ. In column 5, this size effect is picked up by a smaller yet precisely estimated coefficient on GDP shares, $\hat{\gamma}_{\mathrm{GDP}}$. This coefficient is also statistically significant when considering the ban of trade in controlled substances (column 3). There, $\hat{\gamma}_{\mathrm{GDP}}$ implies a 24% acceleration in the mean ratification time relative to the non-strategic counterfactual, which possibly includes the effect of social preferences not explicitly accounted for in this specification.

4.6. *Counterfactual analysis and implications for treaty design*

A strength of the structural econometric approach lies in its suitability for counterfactual analysis. So far, I have used counterfactuals as benchmarks for comparing the magnitude of strategic complementarity across specifications and the contribution of the different γ coefficients to the overall effect.

Here I conduct further counterfactual experiments that shed light on different aspects of treaty design and negotiation strategy. I start by considering a scenario where a coalition of countries ratifies the treaty within 180 days after it was opened for signature. I implement this by assigning to each coalition member a subsidy that, on day 180, makes the country just indifferent between ratifying or not, assuming that the agreement consists of coalition members only.[25] Early ratification by the coalition likely accelerates the subsequent ratification pattern, but the magnitude of this effect depends on the particular coalition under consideration. Moreover, since the indifference condition implies a negative subsidy for some members of certain coalitions, there is scope for side payments among early movers which can bring down the cost of coordinating on early ratification.

Table 3 reports simulation results for coalitions comprised of (1) Member States of the European Union, (2) the countries with the highest CFC consumption levels and a joint share of 90% of the total, (3) all producer countries and (4) all exporters of controlled substances with a share of at least 5% in the sample. For each coalition, I calculate the percentage increment in the average ratification time resulting from early ratification by the coalition compared to the fitted model (with strategic interaction but without early movers). I separately report the increment among non members so as to highlight the externality caused by the coalition. Moreover, I report the total reduction in the ratification time per unit of ϕ, both with and without transfers. For simplicity, I normalize this efficiency measure using the case of EU accession without transfers

25. The earliest ratification in the data occurs on day 197.

TABLE 3
Counterfactual experiment: early ratification by a coalition

Coalition	(1) European Union	(2) Top CFC Consumers	(3) CFC Producers	(4) Top CFC Exporters
Acceleration of mean ratification times in per cent of ratification times				
Among non-members	0.3	0.8	1.9	0.1
Among all countries	1.6	3.9	6.3	0.8
Reduction in ratification time per unit of ϕ (relative to column 1)				
Without transfers	1.0	1.7	2.2	3.1
With transfers	1.0	3.4	4.1	3.6

Notes: All coalition members ratify before 180 have passed since the treaty opened for signature. EU countries: Belgium, Luxembourg, France, Italy, The Netherlands, Denmark, Greece, Spain. Top CFC consumers: U.S., Japan, France, Italy, The Netherlands, China, Canada, Spain, Australia, South Africa, Brazil, Mexico, Switzerland, Austria, Algeria, Denmark. CFC Producers: U.S., Japan, France, Italy, The Netherlands, Spain, Canada, Australia, Greece, China, South Africa, Brazil, Mexico, Argentina, Venezuela, India. Top CFC exporters: Canada, U.S., Japan, Belgium/Lux, France, Italy, The Netherlands.

as the reference. The simulations are based on the specification using export shares in controlled substances, reported in column 3 of Table 2 above.[26]

The results show that all coalitions accelerate ratification by non members, albeit to a different extent. The effect is largest for the coalition of CFC producers, which also causes the largest overall reduction in ratification times. In contrast, the most efficient coalition consists of the top CFC exporters, which accelerates ratification times more than three times more per unit of cost than a coalition of EU-12 countries. However, when the coalition is allowed to share the cost of early action in an efficient manner—by taxing net gainers within the coalition and using the revenue to compensate net losers—then a coalition of CFC producers is the most cost effective negotiation strategy for reducing the aggregate time to ratification.

In another counterfactual experiment I assume, inspired by the Kyoto Protocol, that the U.S. never ratifies and that there is no trade ban on controlled substances. I find that U.S. non participation alone delays the average ratification time of the other countries by 19% and, combined with an abolishment of the trade ban, by 22% compared to the fitted model.

These numbers strongly suggest that early adoption by the U.S., the world's largest producer and consumer of CFCs, was crucial for jump-starting the ratification process, and that the trade ban acted as a powerful catalyst. Further, the experiments on early adoption support the view that securing early participation by (a coalition of) major CFC producers was the right negotiation strategy initially. The finding that some coalition members can compensate others for the incremental cost of ratifying early provides a rationale for why such payments were offered to producers in developing countries under the London Amendments.

5. INTERNATIONAL TRADE AND INVESTMENT AGREEMENTS

The above application has considered international trade relationships between countries only through the lens of enforcing participation in a global environmental agreement. A different though related question concerns the role of strategic complementarity in explaining the rapid

26. Table D.9 of the Online Appendix reproduces the results when simulations are based on the specification using CFC consumption shares, reported in column 5 of Table 2.

proliferation of agreements over trade and investment. This section applies the strategic estimation framework developed above to shed light on this phenomenon, taking explicitly into account the externalities that those treaties create for non-Member States. In particular, I analyse the timing of PTAs between the EU and third countries, as well as the timing of Germany's BITs with more than 130 countries in the world.

5.1. *Preferential trade agreements*

PTAs have been growing rapidly since the late 1950s. By 2014, almost 300 such agreements were in force (IMF, 2015). A PTA is the legal framework by which two or more countries form a union in which goods produced within the union are subject to lower trade barriers than the goods produced outside the union (Panagariya, 1999).

According to theory, a PTA shifts production from inefficient domestic providers to efficient union members (trade creation) while also diverting trade from efficient third-country suppliers towards union members that enjoy preferential treatment (Viner, 1950). The discriminatory features of PTAs and their potential for trade diversion are at the centre of a scholarly debate about the welfare effects of regional trade liberalization (Freund and Ornelas, 2010). While some view PTAs as a threat to deeper multilateral trade liberalization (Bhagwati, 2008), others have argued that the externalities created by PTAs on non-members further reinforce their proliferation, thus causing trade barriers to fall like dominoes (Baldwin, 1993). The theories underlying both views critically depend both on the type and magnitude of trade diversion and the political economy of tariff-setting. The domino theory maintains that exporting industries in non-member countries lobby for participation in the PTA in order to avoid the adverse effects of trade diversion (Baldwin, 1993). In contrast, Grossman and Helpman (1995) consider a model where special interest groups lobby precisely for the most trade-diverting PTAs in order to gain enhanced protection from non-members. The theoretical ambiguity has motivated a sizeable number of empirical studies on the determinants of PTA formation, on their impact on trade and on their interaction with the multilateral trading system (see Freund and Ornelas, 2010, for a comprehensive survey).

One strand of this literature focuses on the economic determinants of PTA formation. Baier and Bergstrand (2004), Egger and Larch (2008), Baldwin and Jaimovich (2012), and Jaimovich (2012) use large dyadic data sets to estimate binary choice models of the likelihood that country pairs form a PTA. Bergstrand *et al.* (2016) estimate a duration model where the year of PTA formation between a country pair is the outcome of interest. To account for externalities in PTA formation, Egger and Larch (2008) estimate a probit model where the latent benefit of a PTA depends on the spatially weighted number of PTAs in place within a 2,000 km radius. They find evidence that a PTA between a country pair ij is more likely to occur if a nearby country pair kl has also formed a PTA. Baldwin and Jaimovich (2012) combine the spatial probit approach with trade-based "contagion" weights that proxy for the magnitude of trade diversion induced by the PTA between k and l on countries i and j. They find that a nearby PTA increases the probability of a PTA formation over and above the proximity effect found by Egger and Larch (2008) if its predicted trade diversion effect is high, which lends empirical support to the domino theory of regionalism.

The key strengths of this literature lie in the exploitation of large data sets and the use of well-known econometric techniques to make inference in the presence of unobserved heterogeneity and cross-sectional dependence. However, the identifying assumptions also imply that countries do not form expectations about PTA formation among their trading partners. Rather, they merely react to the formation of a new agreement after the fact. This assumption can be challenged on the grounds that trade negotiations often drag on for years and receive a lot of media attention. Therefore, I take a different approach and model regional trade integration as a dynamic game played by forward-looking governments capable of anticipating that new PTAs give rise to trade diversion. This is in line with the wide-spread view that governments take into account the strategic effects of trade policy.

5.1.1. PTA formation with the EU as a dynamic game. One of the first and most eminent PTAs is the European Economic Community, which was founded in 1959 by Belgium, France, Germany, Italy, Luxembourg and the Netherlands (henceforth referred to as the EC6 countries). During the post-war period, the EC6 has developed the largest network of PTAs in the world (Ahearn, 2010). The formation of this network can be cast as a dynamic game of treaty formation along the lines of the framework developed above. Countries have perfect information and are free to accede and withdraw from the PTA in any period. Each country maximizes the present discounted value of cumulative per-period payoffs, given in equation (2.2). The private net benefit of not signing a PTA with Europe is normalized to 0. Signing is associated with a private net cost ϕ_i which is gradually offset by the monotonically increasing trend $g(t)$. This accounts for the possibility that the PTA becomes more attractive over time for reasons unrelated to other countries' actions, *e.g.* due to GDP growth in Europe or because of decreasing trading costs. Recall that, from an econometric point-of-view, $g(t)$ is the trend against which the model has to distinguish strategic effects that either accelerate or delay PTA formation.

Strategic interaction may arise if the PTA induces trade diversion. Given two countries that both trade with the EU, the signing of a PTA between country A and the EU diverts some of the trade between the EU and country B towards country A, and hence reduces country B's payoff to not signing a PTA with the EU ($\gamma_0 < 0$). By signing a PTA with the EU, country B benefits from a reversion of the trade diversion effect, and, possibly, from trade creation with the EU. Both effects imply that $\gamma_1 > \gamma_0$ and hence $\gamma > 0$ in the relative payoff function given by equation (3.14). Trade diversion thus creates strategic complementarity and accelerates the formation of PTAs.

The literature suggests that, if country B trades a lot with country A prior to the agreement, it should be more affected by trade diversion than a country C that engages in little or no trade with country A. Following this logic, I weight the impact of A's ratification decision on country B by the share of exports from B to A in B's total exports, as in equation (4.20) above. This is an upper bound on imports that country A may divert from B to the EU after forming the PTA.[27]

Since export flows may be endogenous to PTA formation, I also compute the export weights in equation (4.20) based on exports predicted using a gravity equation.[28] In addition, I compute weights based on the inverse distance between countries i and j,

$$w_{ij} = (\text{population weighted distance between } i \text{ and } j)^{-1} \tag{5.24}$$

and weights that account only for the ratification of contiguous countries,

$$w_{ij} = \frac{\mathbb{1}\{i \text{ and } j \text{ share a common border}\}}{\sum_{k \neq i} \mathbb{1}\{i \text{ and } k \text{ share a common border}\}}. \tag{5.25}$$

Being based on exogenous measures of trading costs, these "geographic" weights are arguably correlated with i's share of trade with country j and, hence, with the expected trade diversion when j enters a PTA with the EU.

27. In addition, the PTA may divert bilateral trade flows between the EU and country B towards country A. This effect should be proportional to pre-existing trade with the EU and depends on trading costs, for which I control in the term ϕ by including the distance between a country and the nearest EC6 country.

28. See Section E of the Online Appendix for more details.

5.1.2. Data. Data on PTAs between Europe and the rest of the world were downloaded from the internet portal of the DG Trade of the EU Commission.[29] I consider the entire history of PTAs between the EC6 countries and third countries. The process starts with the establishment of the European Economic Community in the Treaty of Rome. The treaty was signed on 25 March 1957, notified to GATT on 24 April 1957, and it entered into force on 1 January 1958. The date relevant for my analysis is the date of notification, which is the earlier of the two dates that are available in the database.

Data on country characteristics and bilateral commodity trade are taken from the GeoDist and Gravity databases provided by CEPII (Head *et al.*, 2010; Mayer and Zignago, 2011). The set of covariates shifting ϕ_i follows the empirical literature which emphasizes the proximity of the trading partners in geographic, cultural and economic terms. Specifically, I include a country's log per capita income, log population, the minimum distance to the EC6 countries (based on the population-weighted distance between metro areas) and a measure of cultural proximity between a country and the EC6 states.[30] Because the fall of the iron curtain fundamentally changed the geopolitical landscape, I also include a dummy for the fall of the Berlin wall on 9 November 1989. Panels A and B of Table E.1 in the Online Appendix summarize, respectively, the full data set of 69 ratification events and a smaller data set consisting only of the fourty ratification events following the fall of the iron curtain.

5.1.3. Results. Table 4 summarizes the estimates for the strategic parameters of the model of EU trade agreements. Column 1 reports the results for the model of global interaction effects (unrestricted γ and symmetric weights). The estimated $\hat{\gamma}$ is negative and statistically significant, suggesting that PTA formation is a strategic substitute. Panel B of the table reports that this effect delayed the formation of PTAs with the EU by 2,254 days or by 18% compared to a non-strategic counterfactual.

This finding is robust to dropping the first twenty-nine PTAs from the sample and performing the estimation only on the fourty PTAs that were signed after the fall of the iron curtain (cf. Table E.2 in the Online Appendix). Although the model with symmetric spillovers rejects strategic complementarity, I re-estimate the model with asymmetric spillovers and non-negativity constraints on γ. This yields comparatively small and statistically insignificant point estimates for both γ's (columns 2–5 of Table 4). As shown in Panel B, the point estimates imply 0% acceleration in ratification time compared to the non-strategic counterfactual, *i.e.* the non-negativity constraints are binding. In that sense, the estimation framework proposed here is robust to predicating the estimation on an incorrect assumption about the nature of strategic interaction.

The estimation results do not support the domino theory of trade liberalization. Since this effect hinges on the presence of trade diversion, a reason for this could be that "Europe's PTAs for the most part have not liberalized substantial amounts of trade" (Ahearn, 2010, p. 28). This is in line with the argument that Europe's PTAs overemphasize the promotion of EU standards as opposed to the opening of closed sectors to trade (Bhagwati, 2008).

29. Available Online at http://ec.europa.eu/trade/creating-opportunities/bilateral-relations/ (last accessed on January 15, 2016).

30. From the raw data I obtain dummy variables for (i) a common colonizer after 1945, (ii) a colonial relationship after 1945, (iii) a colonial relationship ever, (iv) a colonial relationship currently, (v) a common official language, and (vi) a common language spoken by >9% in both countries. After taking the maximum value (across the EC6 countries) for each indicator, I add up the six indicators and standardize the resulting variable by its mean and standard deviation.

TABLE 4
PTAs with the EU: estimation results

	(1)	(2)	(3)	(4)	(5)
	Global Interaction		Sources of Complementarity		
Weights w_{ij}	Equal	Distance	Contiguity	Exports	
				actual	predicted
A. Coefficient estimates γ					
γ	−61,140	22.140	0.003	474.960	30.624
	(5,344.8)	(49.563)	(0.008)	(755.660)	(87.020)
γ_{GDP}		1.302	1.608	2.281	3.355
		(6.554)	(2.669)	(4.533)	(6.690)
B. Delay in mean ratification time w.r.t. non-strategic counterfactual					
γ	−2,254 (−18%)	0	0	0	0
γ_{GDP}		0	0	0	0

Notes: $N=69$. Bootstrapped standard errors are in parentheses. The full set of parameter estimates is reported in Table E.2 in the Online Appendix.

5.2. *Bilateral investment treaties*

A BIT is an intergovernmental agreement that provides legal stability and protection of FDI. By signing a BIT, two countries agree on a set of conditions concerning the admission and treatment of FDI from the home country in the host country. The typical BIT grants extensive rights to foreign investors, including the right to bring investment disputes before international arbitration venues such as the World Bank's International Centre for the Settlement of Investment Disputes (ICSID) or the UN Commission on International Trade Law (UNCITRAL).

Since the signing of the first BIT between Germany and Pakistan in 1959, the number of BITs in the world has been growing, at first moderately until the mid-1980s, then at a rate of more than a hundred new treaties per year throughout the 1990s. At present, there are 2,922 BITs on record, of which 2,240 are in force (UNCTAD, 2015). While the traditional BIT involves a developed home country and a developing host country, developing countries have increasingly been forging BITs among themselves.

Despite their focus on FDI, BITs thus far have not received much attention by academic economists. There is, however, a political science literature on BITs which coincides in that "BITs have become the most powerful international legal mechanism for the encouragement and governance of FDI" (Elkins *et al.*, 2006).[31] The obvious question of whether the enactment of a BIT causes an increase in FDI has been addressed in a series of empirical studies (Hallward-Driemeier, 2003; Tobin and Rose-Ackerman, 2003; Kerner, 2009; Egger and Merlo, 2012). Rather than offering a conclusive answer to this question, this strand of the literature highlights the challenging endogeneity issues facing researchers who tackle this issue (Aisbett, 2007). Looking beyond the mere existence of a BIT, empirical research suggests that the treaty's ability to attract FDI hinges not so much on the stringency of dispute-settlement provisions granted to the investor but rather on the good behaviour of the host government subsequent to signing the BIT (Allee and Peinhardt, 2011; Berger *et al.*, 2011, 2013). Furthermore, there is empirical

31. Much of this literature deals with the paradox that the "Hull rule"—a concept of customary international law which prohibited expropriations and provided for prompt and adequate compensation if they occurred—gave way to a regime of *bilateral* treaties that implemented even stricter standards of investment protection, rather than to a *multilateral* investment agreement (*e.g.* Guzman, 1998).

evidence that the existence of a BIT is positively associated with debt accumulation (Mina, 2013) and subsequent PTA formation (Tobin and Busch, 2010).

Another strand of the literature argues that international competition for FDI among potential host countries has been driving the widespread adoption of BITs (Guzman, 1998). To test this hypothesis, which will be explained in more detail below, Elkins *et al.* (2006) combine data on the diffusion of bilateral BITs signed between 1962 and 2000 with measures of economic competition among host countries. Using a Cox duration model for home/host country dyads, they find that potential hosts are more likely to sign BITs when their competitors have done so. Neumayer and Plümper (2010) show that, rather than a capital-importing country being influenced by the total number of BITs signed by other capital importers, as modelled in Elkins *et al.* (2006), a capital-importing country is more likely to sign a BIT with a capital exporter only if other competing capital importers have signed BITs with this very same capital exporter.

5.2.1. Bilateral investment treaties as a dynamic game. While Guzman (1998) and Elkins *et al.* (2006) do not present formal models, it is straightforward to cast their "competition-for-capital" hypothesis as a dynamic game using the framework developed above. The game is played by N potential host countries that compete for a fixed pool of FDI from the "home" country. All countries decide in each period whether they sign a BIT with the home country. The home country has all the bargaining power over the core terms as host countries compete with other potential hosts. Therefore, the treaty is akin to a take-it-or-leave-it offer, the terms of which are always attractive for the home country.

Elkins *et al.* (2006) argue that this is an accurate representation of the classical BIT where capital flows from a developed country to a host country in the developing world. By taking the diffuse, multilateral commitment to customary law to the bilateral level, a BIT allows the host countries to signal their intent of contracting in good faith. The signal is credible because it generates diplomatic, sovereignty, arbitration, and reputation costs in its observance and violation. The treaty thus increases expected returns for investors and hence their propensity to invest in the host country. At the same time, the BIT "gives host governments a competitive edge in attracting capital" over rival countries, especially if there are otherwise doubts about their willingness to enforce contracts fairly (Elkins *et al.*, 2006, p. 823).

Suppose the per-period payoff of a host country is given by equation (2.2), where the private net benefit of not signing the BIT is normalized to zero. As more of its rivals sign a BIT with the home country, country i sees its prospects of receiving FDI diminished, *i.e.* $\gamma_0 < 0$. In contrast, by signing the BIT the country can (at least partially) mitigate this negative externality of BITs signed by its rivals, *i.e.* $\gamma_1 > \gamma_0$. In addition, there may be a benefit of being the only host country to have signed a BIT. On the other hand, signing the BIT is costly in that the host country gives up sovereign rights in the treatment of FDI, and because international arbitration raises the costs of non-compliance.[32] Elkins et al. assume that this benefit is not large enough to drive a host country into signing a BIT in the absence of other BITs, *i.e.* the private net cost of signing ϕ_i is strictly positive. I specify the relative payoff function to BIT ratification as in equation (3.14) above. By decreasing the private net cost the time trend $g(t)$ nests the hypothesis that the proliferation of BITs is propelled by non-strategic drivers, such as (exogenous) growth in the overall pool of FDI available from the home country. Competition for capital implies that $\gamma = \gamma_1 - \gamma_0 > 0$. This

32. The typical BIT provides more substantive protections than customary law, *e.g.* national treatment and most-favoured-nation treatment of FDI in the host country, protected contractual rights, a guarantee to transfer profits in hard currency to the home country and prohibited or restricted use of performance requirements (Elkins *et al.*, 2006).

TABLE 5
Germany's BITs

Weights w_{ij}	(1) Equal	(2) Distance	(3) Contiguity	(4) Exports	(5)
				actual	predicted
A. Coefficient estimates					
γ	4,353	31,171	2,073	5,393	96.099
	(1,240)	(8,749)	(470)	(1,925)	(76.638)
γ_{GDP}		488.570	17,926	810.310	1,812.800
		(145.010)	(4,284.100)	(331.73)	(587.300)
B. Acceleration of mean ratification time w.r.t. non-strategic counterfactual					
Δ in days					
γ	2,522	2,767	1,694	1,217	1
γ_{GDP}		1,334	496	690	3,549
both		2,914	2,133	1,286	3,552
Δ in percentage					
γ	27	22	14	10	0
γ_{GDP}		11	4	6	29
both		24	17	11	29

Notes: $N = 127$ countries. All specifications assume coordination. Bootstrapped standard errors are in parentheses. The full set of parameter estimates is reported in Appendix Table E.3.

completes the representation of the "competitive dynamic among potential hosts" depicted in Elkins *et al.* (2006) as a dynamic game of BIT formation with strategic complementarity.

As in the previous applications, additional information can help with the identification of the complementarity parameter γ. Elkins *et al.* (2006) argue that potential host countries—analogous to differentiated products—are close substitutes for FDI if they have similar attributes. The intensity of competition between two countries i and j should thus be reflected by their proximity in attribute space w_{ij}. As in the model of PTA formation, I focus on geographic attributes which are exogenous to treaty formation by specifying weights based on distance shown in equation (5.24) and contiguity equation (5.25). In the BIT application, however, these attributes are not used as proxies for trade flows but rather as direct shifters of the profitability of FDI. To allow comparisons between the two models, I also consider trade weights equation (4.20) based on both actual and predicted trade flows.

5.2.2. Data. Data on bilateral investment agreements are available from the International Investment Agreements Navigator website maintained by UNCTAD (2015). Germany signed the first BIT on 25 November 1959 with Pakistan and has since signed BITs with 134 countries, more than any other country in the world. I augment data on Germany's BITs with data on income per capita, population, distance to Germany (all in logs), a counting measure of cultural ties with Germany and a dummy variable for BITs signed after the fall of the Berlin Wall in November 1989. I construct these data in the same fashion as described above in the context of the PTA model. The final data set comprises 127 BITs and is summarized in Table E.1 in the Online Appendix.

5.2.3. Results. Table 5 summarizes the results for the strategic model of Germany's bilateral investment agreements. The first column suggests that strategic complementarity played a statistically and economically significant role in explaining these treaties, accelerating BIT formation with the average country in the sample by almost 7 years or 27% compared to the non-strategic counterfactual. The results reported in columns 2–3 can be interpreted as supporting the hypothesis that host countries competed for German FDI. In particular, when the interaction is limited to countries that are close (contiguous), and size effects are controlled for by a separate parameter γ_{GDP} on GDP, the competition effect accelerates the mean time to BIT signing by 22% (14%) relative to the non-strategic counterfactual. Columns 4–5 show that export shares matter less, and—it seems—only to the extent that they proxy for geographic distance. When predicted export shares based on gravity models are used to measure the intensity of competition, no acceleration of BIT formation is measured.

These findings are consistent with the competition-for-capital hypothesis in the context of the typical match between a capital-exporting industrialized country and developing host countries. Many of the more recent BITs defy this stereotype as they are formed by developing countries only. Competition for capital may take a different shape in this context. In fact, when I fit a similar model to data on 102 BITs formed by China since 1982, I obtain positive yet statistically insignificant point estimates of γ in all but one specification (see Table E.4 in the Online Appendix). This suggests that there are important differences in the patterns of strategic interaction engendered between countries, even when the treaty contents are very seemingly similar.

6. CONCLUSION

This article has developed a general framework for empirically analysing the formation of international agreements. Casting treaty ratification as a dynamic game, the framework nests strategic and non-strategic mechanisms that drive countries to participate in an agreement, and is amenable to econometric estimation using a single history of ratification dates. The principal benefits are twofold. Unlike conventional duration models, my framework explicitly takes into account strategic, forward-looking behaviour on the part of a government that decides on participation in an international treaty. Moreover, by estimating the behavioural parameters of a structural model, the framework can be used to conduct counterfactual experiments that explore specific channels of interaction.

My framework is suitable for analysing a broad class of international treaties. I demonstrate this in applications to a global environmental agreement, to preferential trade agreements and to bilateral investment treaties. Over the past decades, international treaties in these policy domains have been proliferating. While theoretical research has emphasized the scope for strategic behaviour in each of these domains, it is far from clear to what extent strategic considerations have contributed to this proliferation. The method proposed in this article provides applied researchers with a model-based tool for empirically analysing this aspect of international treaty-making.

Empirical research on international agreements is subject to the fundamental limitation that there are just about 200 countries in the world. The data-poor environment means that strong assumptions need to be made before deeper insights can be gained from observed treaties. For instance, analysing pooled data from different treaties requires an assumption that the distribution of costs and benefits is identical across treaties. This article takes a different road by trying to learn about strategic interaction from a single cross-section of ratification times. Inevitably, this approach must be predicated on assumptions, some of which are strong. While some discussion has already been provided when these assumptions were introduced, I expand on this discussion here for the benefit of readers who are interested in building and extending on this framework in their research.

A fundamental pair of assumptions is that the variation in the time to treaty ratification contains information about payoffs, and that it is not driven by common shocks to these payoffs. These assumptions are not testable per se. Econometrically, controlling for common shocks in this data environment is a formidable task. However, developing a framework that allows for time-varying covariates could help to mitigate concerns about common shocks confounding the estimates of γ.

Another set of assumptions pertains to the specification of the model and functional forms of time and country effects. Critical such assumptions are that the time effect enters the payoff function additively and is common across countries. This is needed for identification because otherwise ratification events could be clustered also in the absence of strategic complementarity, simply because the relative payoff functions for two countries cross the treshold at the same time. Moreover, this would break the link between the order of ratification and the net cost of ratification, which is needed for the identification of ϕ. It is possible to relax these assumptions in a hypothetical application where the time effect is known (*e.g.* calibrated to a country's GDP growth), though this may in itself require some assumptions. Apart from that, there may be a small set of treaties for which pooling data is a defendable choice. In that case it is possible to identify some of the objects in the payoff function under weaker assumptions on functional form, *e.g.* by looking at order statistics across treaties.

Further assumptions enter via modelling choices such as common knowledge and perfect foresight on behalf of governments. This may be more plausible than assuming the exact opposite (*i.e.* no anticipation of future ratification choices), but it is clear that accounting for some of the uncertainty surrounding treaty formation would add realism to the theoretical model.

I model treaty formation as binary choice problem. A worthwhile extension would be to explicitly account for the level of participation, such as the reduction target for pollution emissions or tariff rates. Related to this, extending the model to account for a richer set of state variables would be of great value in applications such as agreements over stock pollutants.

These and other extensions are left as a topic for future research.

APPENDIX

Proof of Theorem 1. The action space $A=\prod_{i=1}^{N} A_i$ is a complete lattice because it is the direct product of N finite chains, A_i. By assumption $\gamma \geq 0$, the payoff function $\pi_i(a,t)$ is supermodular on A for each $t \geq 0, i \in I$ (see Corollary 2.6.1 in Topkis, 1998) and hence $\Gamma(t)$ is a supermodular game $\forall t \in T$. Under these conditions, Topkis (1979) and Vives (1990) have shown that the set of Nash equilibria is non-empty and contains a greatest and a least element. Milgrom and Roberts (1990) have proven that these elements are pure-strategy Nash equilibria and that they are non-decreasing functions of the parameter t. Moreover, they have shown that the largest Nash equilibrium is the most preferred one for all players. ‖

Proof of Corollary 1. In a dominant strategy equilibrium, the best response correspondence $R(a_{-i},t) \equiv \arg\max_{a_i \in A_i} \pi_i(a_i, a_{-i}, t)$ is constant on A. Suppose first that $R(\cdot,t)$ is not constant. Then there exists some $i \in I$ and $a_{-i}, a'_{-i} \in A_{-i}$ such that $a_{-i} \leq a'_{-i}$ and $R_i(a'_{-i},t) \neq R_i(a_{-i},t)$. From assumption $\gamma=0$ we have that

$$\pi_i(R_i(a'_{-i},t),a'_{-i},t)-\pi_i(R_i(a_{-i},t),a'_{-i},t)=\pi_i(R_i(a'_{-i},t),a_{-i},t)-\pi_i(R_i(a_{-i},t),a_{-i},t). \tag{A.26}$$

By the definition of the best response, the LHS of this expression is non-negative and the RHS is non-positive. Hence both sides of the equation must be equal to zero. Since ties for different actions are ruled out, the equation can only hold if $R_i(a'_{-i},t)=R_i(a_{-i},t)$, a contradiction. Thus, each player has a strictly dominant action and the profile of dominant actions constitutes the unique Nash equilibrium in period t. ‖

Proof of Theorem 2. I first use backward induction and the dominance property to prove that the SRP induces the largest Nash equilibrium in every stage of the game. Next, I establish that the profile s^* defined above selects the largest Nash equilibrium of the stage game in every period $t \in T$.

Step 1: SRP equilibrium induces the largest Nash equilibrium in every stage game. Denote by t_N^0 the first period in which cooperation is a dominant strategy for all players. Full cooperation is the unique NE of the stage game $\Gamma(t_N^0)$ and, by the monotonicity of equilibrium stated in Theorem 1, in all subsequent periods. Indefinite repetition of this profile constitutes a subgame-perfect Nash equilibrium of the continuation game $G(t_N^0)$. The equilibrium is WRP because it is not dominated by any of its continuation equilibria. Since the action profile $a^N=\mathbf{1}$ dominates every other profile in the stage game, there can be no other WRP equilibrium that dominates indefinite cooperation. Hence full cooperation is the unique SRP equilibrium of the continuation game $G(t_N^0)$.

It follows that there is no credible "punishment" that could be inflicted on a player who deviates in period $t_N^0 - 1$ and hence that any subgame-perfect equilibrium induces a Nash equilibrium of the stage game at node $t_N^0 - 1$. Backward induction yields that a stage-game Nash equilibrium must be played in all previous periods, too. Moreover, the refinement of strong renegotiation-proofness requires that none of these stage-game equilibria be Pareto-dominated, for any such WRP profile $\tilde{t}$ would be dominated by another WRP profile $\tilde{s}'$ that induces the largest Nash equilibrium in stage t' and is otherwise identical to $\tilde{s}$. Hence the SRP equilibrium induces the largest Nash equilibrium in every stage game.

Step 2: The profile s^* induces the largest Nash equilibrium in every stage game. By the monotonicity of the Nash equilibrium vector in t, a player who cooperates in some period t' will not revert her decision in any later period $t'' > t'$. For instance, the last player $j_{(N)}$ to cooperate in period $t^*_{(N)}$ strictly prefers to defect in period $t^*_{(N)} - 1$. Thus $j_{(N)}$ cannot be part of the largest (and hence: any) Nash equilibrium in the periods $t \leq t^*_{(N)} - 1$. Solving for the equilibrium path hence boils down to finding, recursively for each player $i \in N$, the period in which i cooperates for the first time in the largest Nash equilibrium.

The principle of induction is invoked to prove that, for each $k = 1, \ldots, N-1$ the strategy profile s^* induces the largest Nash equilibrium in the periods from $t^*_{(N-k)}$ until (but not including) $t^*_{(N-k+1)}$, *i.e.* until $\max[t^*_{(N-k)}, t^*_{(N-k+1)} - 1]$. Consider the base case $k = 1$. Suppose first that $t_{(N-1)} \geq t^*_{(N)}$. For any such $t_{(N-1)}$ the strategy profile s^* induces full cooperation, *i.e.* the largest Nash equilibrium of the stage game from period $t^*_{(N)}$ onwards. Next, suppose that $t_{(N-1)} < t^*_{(N)}$ and the algorithm sets $t^*_{(N-1)} = t_{(N-1)}$ according to equation (2.8). The definition of $t_{(N-1)}$ implies that cooperation by the players in K_{N-1} is a Nash equilibrium of the game $\Gamma(t_{(N-1)})$ and that full cooperation is not. By monotonicity, this is true for all $t_{(N-1)} \leq t < t^*_{(N)}$.

The inductive hypothesis maintains that, for some k such that $1 \leq k \leq N-1$, the strategy profile s^* induces the largest Nash equilibrium in the stage game from period $t^*_{(N-k)}$ through $\max\left[t^*_{(N-k)}, t^*_{(N-(k-1))} - 1\right]$. The inductive step is to prove that the profile s^* induces the largest NE in all periods from $t^*_{(N-(k+1))}$ until $\max\left[t^*_{(N-(k+1))}, t^*_{(N-k)} - 1\right]$. For $t_{(N-(k+1))} \geq t^*_{(N-k)}$ the algorithm assigns $t^*_{(N-(k+1))} = t^*_{(N-k)}$ and the inductive step follows trivially from the inductive hypothesis.

Suppose now that $t_{(N-k-1)} < t^*_{(N-k)}$. Then $t^*_{(N-k-1)} = t_{(N-k-1)}$ and s^* induces the profile a^{N-k-1} with the set of cooperators $K_{N-k-1} = K_{N-k} \setminus \{j_{(N-k)}\}$. From equations (2.4) and (2.6), the profile a^{N-k-1} is a Nash equilibrium of the game in period $t_{(N-k-1)}$. For any larger Nash equilibrium $\tilde{a} \geq a^{N-k-1}$ in the stage game $\Gamma(t^*_{(N-k)} - 1)$ monotonicity requires that $\tilde{a} \leq s^*(t^*_{(N-k)})$. Without loss of generality assume that $s^*(t^*_{(N-k)}) = a^{N-k+l}$, *i.e.* $N-k+l$ players cooperate in period $t^*_{(N-k)}$. The "cluster size" $l \in \{0, \ldots, k\}$ is given by the value of l that satisfies

$$t^*_{(N-k)} = t_{(N-k+l)}. \tag{A.27}$$

To see that $\tilde{a} \neq a^{N-k+l}$ let $j = N-k+l$ and write

$$\Delta\pi_{(j)}(a^j_{-j}, t^*_{(N-k)} - 1) = \Delta\pi_{(j)}(a^j_{-j}, t_{(j)} - 1) < 0$$

where we have used equations (2.7) and (A.27). For $\tilde{a} \leq a^{N-k+l}$ player $j_{(N-k+l)}$'s deviation incentive is equally strong ($\gamma = 0$) or stronger ($\gamma > 0$). Thus any profile in which $j_{(N-k+l)}$ cooperates cannot be a Nash equilibrium of the stage game $\Gamma(t^*_{(N-k)} - 1)$.

For $l > 0$, the logic of this argument applies to all other players who cooperate under a^{N-k+l} but not under $a^N - k - 1$. For all $j = N-k+l-1, \ldots, N-k$ it is true that

$$\Delta\pi_{(j)}(a^j_{-j}, t^*_{(N-k)} - 1) \leq \Delta\pi_{(j)}(a^j_{-j}, t_{(j)} - 1) < 0$$

where the first inequality follows from Assumption 3 and the fact that $t^*_{(N-k)} < t_{(j)}$ and the second inequality follows from equation (2.7).

Thus, it must be the case that $\tilde{a} = a^{N-k-1}$ and s^* induces the largest Nash equilibrium in period $t^*_{(N-k)} - 1$. Monotonicity implies that this is true for all periods t such that $t_{(N-k-1)} \leq t < t^*_{(N-k)}$. By the principle of induction, the claim must be true for all $k = 1, \ldots, N-1$.

From equation (2.7) recall that player $j_{(1)}$ has no incentive to join the coalition at any time before $t^*_{(1)}$. Therefore, the largest Nash equilibrium in periods $t = 0, \ldots, t^*_{(1)} - 1$ has no player cooperate. This concludes the proof that s^* induces the largest Nash equilibrium in every stage game. From step 1, this profile constitutes the unique SRP equilibrium of the game. ‖

Proof of Theorem 3. Existence is proven by Berry (1992) in the context of a static entry game. The proof is by arranging countries in increasing order of ϕ_i and finding the equilibrium number of treaty members m that satisfies

$$-\phi_m + g(t) + \gamma(m-1) \geq 0 > -\phi_{m+1} + g(t) + \gamma m. \tag{A.28}$$

Since $\Delta\pi_i(a_{-i}, t)$ is strictly decreasing in a_{-i} the equilibrium number of treaty participants m is unique. Multiplicity arises when ϕ_m and ϕ_{m+1} are close so that

$$-\phi_{m+1} + g(t) + \gamma(m-1) \geq 0 > -\phi_m + g(t) + \gamma m, \tag{A.29}$$

also holds.

Suppose that m is the Nash equilibrium number of signatories in period t, and that the Nash equilibrium is not monotone. Then there must be a period $t' > t$ and a k such that $-\phi_k + g(t) + \gamma(m-1) \geq 0$ and $-\phi_k + g(t') + \gamma(m-1) < 0$. This implies $-\phi_k + g(t') > -\phi_k + g(t)$, which contradicts Assumption 3. ||

Proof of Theorem 4. Monotonicity implies that full participation is the unique Nash equilibrium of the stage game from period $t_{(N)}$ onwards. Since no punishment of deviations is possible beyond this point, a stage-game Nash equilibrium must be played in all periods prior to $t_{(N)}$. From Theorem 3, a stage game equilibrium in pure strategies exists and is unique up to the identity of players. When multiple equilibria arise in the stage game, this leads to multiple SRP equilibria as the stage-game equilibria are not ranked. While it is clear from the proof of Theorem 2 that the algorithm used to compute equilibrium selects the largest Nash equilibrium in each stage game, this does not resolve the multiplicity issue because all equilibria have the same number of treaty members. Consider a period t with two equilibria that both have m Member States, as described by equations (A.28) and (A.29) above. Suppose that the algorithm has selected the equilibrium where country $(m+1)$ is a member and not country (m), as in equation (A.29). This means that in some period $t' > t$, the algorithm dropped country (m) from the set of treaty members but not $(m+1)$. From equation (2.6), this only occurs if $-\phi_m + g(t') < -\phi_{m+1} + g(t)$. However, this contradicts $-\phi_m + g(t') \geq -\phi_{m+1} + g(t)$ which is true because we have ordered countries in increasing order of their net cost ϕ. The algorithm circumvents this by always picking the country with a higher $\Delta\pi$ for given m and t, *i.e.* lower cost ϕ. This is precisely how the ambiguity is resolved by the order of moves assumed in the theorem. ||

Remark. As the length of the time periods goes to zero, the order of moves assumption is no longer needed. From Theorem 5, country m will ratify at time $\tilde{t_{(m)}} = g^{-1}[\phi_m - \gamma(m-1)]$ which is strictly earlier than $\tilde{t}_{(m+1)} = g^{-1}[\phi_{m+1} - \gamma(m-1)]$ given Assumption 3 and the definition of $\tilde{t}_{(m)}$ in equation (3.10).

Proof of Corollary 2. Under assumption $\gamma = 0$ equation (2.4) is equivalent to $\Delta\pi_i(0,t) \geq 0$ and condition (2.7) is equivalent to $\Delta\pi_{(m)}(0, t_{(m)} - 1) < 0$. From this we have that candidate equilibrium times satisfy $\Delta\pi_{(m)}(0, t_{(m)} - 1) < 0 \leq \Delta\pi_{(m)}(0, t_{(m)})$ which coincides with the definition of t_i^0 in equation (2.9). To see that $t_{(m)}^*$ always equals $t_{(m)}$, suppose that there is some $l > m$ such that $t_{(l)} < t_{(m)}$. Hence we have that

$$\pi_{(m)}(0, t_{(l)}) \leq \pi_{(m)}(0, t_{(m)-1}) < 0$$

where the first inequality follows from Assumption 3 and the fact that $t_{(l)} < t_{(m)}$. The second inequality follows from equation (2.7). By monotonicity, $a^m \leq a^l$ and equation (2.4) implies that

$$\pi_{(m)}(a^l, t_{(l)}) = \pi_{(m)}(0, t_{(l)}) \geq 0$$

– a contradiction. Hence it must be the case that $t_{(1)} \leq t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(N)}$ which completes the proof. ||

Proof of Theorem 5. Suppose that the SRP equilibrium ratification time of player i in game G is given by $t_i^* = \lceil \tilde{t}_i^* \rceil$. Note that G is equivalent to ${}^0G$. Hence, the equilibrium ratification time of player i in the game ${}^kG$ can be obtained by simply relabelling the decision nodes of the game. This gives ${}^kt_i^* = \lceil 2^k \tilde{t}_i^* \rceil$. The sequence $x_k \equiv \frac{\lceil 2^k \tilde{t}_i^* \rceil}{2^k}$ converges to $\tilde{t}_i^*$ because, for any $\epsilon > 0$, there exists an integer N_ϵ such that $x_k - \tilde{t}_i^* < \epsilon$ for all $k \geq N_\epsilon$. To see this, notice that, by definition of the ceiling function,

$$\frac{\lceil 2^k \tilde{t}_i^* \rceil}{2^k} - \tilde{t}_i^* < \underbrace{\frac{2^k \tilde{t}_i^* + 1}{2^k} - \tilde{t}_i^*}_{=2^{-k}},$$

and let $N_\epsilon = \left\lceil -\frac{\ln\epsilon}{\ln 2} \right\rceil$. By the same token, $x_{k'} - \tilde{t}_i^* < 2^{-k'} < 2^{-k}$ for all $k' > k$. ||

Proof of Theorem 6. Consider a pair of heterogeneous countries i and j with (limit) equilibrium ratification times $\tilde{t}_i^*, \tilde{t}_j^*$, respectively. From Theorem 5, ratification times converge to $\tilde{t}_i^*$ and $\tilde{t}_j^*$ as the grid length goes to 0. In the limit, clustering occurs if and only if $\tilde{t}_i^* = \tilde{t}_j^*$. To see this, consider the relative payoff to cooperation in the absence of strategic complementarity, given by $\Delta\pi_i(a_{-i}, t) = -\phi_i + g(t)$. The equilibrium conditions (3.13) solve for $\tilde{t}_l^0 = g^{-1}(\phi_l)$ where Assumption 3 has been invoked to invert g. Therefore, $\tilde{t}_i^0 = \tilde{t}_j^0$ is equivalent to $g^{-1}(\phi_i) = g^{-1}(\phi_j)$. Since ϕ is a continuous random variable and g^{-1} is strictly monotonic, this event has probability zero. In contrast, if payoff functions exhibit increasing differences ($\gamma > 0$) then it follows from the recursive definition of $\tilde{t}_{(m)}^*$ in equation (3.12) that limit ratification times can be identical even among asymmetric countries. ||

Acknowledgments. This article is based on the first two chapters of my Ph.D. dissertation. I am grateful to my committee members Pat Bayer, Steve Berry, and Chris Timmins for their advice and support. I would like to thank

Marco Ottaviani (the editor) and four anonymous referees for comments that have substantially improved the paper. I have benefited from discussions with Scott Barrett, Dirk Bergemann, Antonio Cabrales, Klaus Desmet, Hanming Fang, Geoff Heal, Nat Keohane, Gerard Llobet, Pei-Yu Lo, Pedro Mira, Marc Möller, Juan Pablo Rincón, Bernard Salanié, François Salanié, and Philipp Schmidt-Dengler. Seminar participants at Alicante, Berkeley, Colby, Columbia, CORE, CSIC, HECER, Hunter, IESE, Maastricht, NYU, Toulouse, and Yale have provided helpful comments. All errors are my own. Jonatan García Salinas provided outstanding research assistance. Oona Hathaway graciously provided me with her data set on human rights treaties. I gratefully acknowledge receipt of an ERP Fellowship from the German National Academic Foundation, a Fisher Dissertation Fellowship from Resources for the Future, and a Georg Walter Leitner Grant in International and Comparative Political Economy from Yale University. Research funded in part by the Spanish Ministry for Science and Innovation (SEJ2007-62908), and by the Spanish Ministry of Economy and Competitiveness (ECO2012-31358, RYC-2013-12492).

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

AAKVIK, A. and TJØTTA, S. (2011), "Do Collective Actions Clear Common Air? The Effect of International Environmental Protocols on Sulphur Emissions", *European Journal of Political Economy*, **27**, 343–351.

ACKERBERG, D. A. and GOWRISANKARAN, G. (2006), "Quantifying Equilibrium Network Externalities in the ACH Banking Industry", *The RAND Journal of Economics*, **37**, 738–761.

AGUIRREGABIRIA, V. and MIRA, P. (2010), "Dynamic Discrete Choice Structural Models: A Survey", *Journal of Econometrics*, **156**, 38–67.

AHEARN, R. J. (2010), "Europe's Preferential Trade Agreements: Status, Content, and Implications", DTIC Document.

AISBETT, E. (2007), "Bilateral Investment Treaties and Foreign Direct Investment: Correlation Versus Causation" (MPRA Paper No. 2255, University Library of Munich, Germany).

ALLEE, T. and PEINHARDT, C. (2011), "Contingent Credibility: The Impact of Investment Treaty Violations on Foreign Direct Investment. *International Organization*, **65**, 401–432.

ASHEIM, G. and HOLTSMARK, B. (2009), "Renegotiation-proof Climate Agreements with Full Participation: Conditions for Pareto-Efficiency", *Environmental and Resource Economics*, **43**, 519–533.

AUFFHAMMER, M., MORZUCH, B. and STRANLUND, J. (2005), "Production of Chlorofluorocarbons in Anticipation of the Montreal Protocol", *Environmental and Resource Economics*, **30**, 377–391.

BAIER, S. L. and BERGSTRAND, J. H. (2004), "Economic Determinants of Free Trade Agreements. *Journal of International Economics*, **64**, 29–63.

BALDWIN, R. (1993), "A Domino Theory of Regionalism", (Working Paper No. 4465, National Bureau of Economic Research).

BALDWIN, R. and JAIMOVICH, D. (2012), "Are Free Trade Agreements Contagious?", *Journal of International Economics*, **88**, 1–16.

BARRETT, S. (1994), "Self-enforcing International Environmental Agreements", *Oxford Economic Papers*, **46**, 878–94.

BARRETT, S. (1997a), "Heterogeneous International Environmental Agreements", In C. Carraro, (ed.) *International Environmental Negotiations. Strategic Policy Issues*, (Cheltenham, UK: Edward Elgar) 9–25.

BARRETT, S. (1997b), "The Strategy of Trade Sanctions in International Environmental Agreements", *Resource and Energy Economics*, **19**, 345–361.

BARRETT, S. (2003), *Environment and Statecraft* (New York: Oxford University Press).

BATTAGLINI, M. and HARSTAD, B. (2015). "Participation and Duration of Environmental Agreements", *Journal of Political Economy*, doi:10.1086/684478.

BENEDICK, R. E. (1998), *Ozone Diplomacy : New Directions in Safeguarding the Planet*. enlarged ed. (Cambridge, MA: Harvard University Press),

BERGER, A., BUSSE, M., NUNNENKAMP, P. and ROY, M. (2011), "More Stringent BITs, Less Ambiguous Effects on FDI? not a bit!", *Economics Letters*, **112**, 270–272.

BERGER, A., BUSSE, M., NUNNENKAMP, P. and ROY, M. (2013), "Do Trade and Investment Agreements Lead to More FDI? Accounting for Key Provisions Inside the Black Box", *International Economics and Economic Policy*, **10**, 247–275.

BERGSTRAND, J. H., EGGER, P. and LARCH, M. (2016), "Economic Determinants of the Timing of Preferential Trade Agreement Formations and Enlargements", *Economic Inquiry*, **54**, 315–341.

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.

BERON, K. J., MURDOCH, J. C. and VIJVERBERG, W. P. (2003), "Why Cooperate? Public Goods, Economic Power, and the Montreal Protocol", *Review of Economics and Statistics*, **85**, 286–97.

BERRY, S. (1992), "Estimation of a Model of Entry in the Airline Industry", *Econometrica*, **60**, 889–917.

BHAGWATI, J. (2008), *Termites in the System: How Preferential Trade Agreements Undermine Free Trade*. (Oxford, UK: Oxford University Press).

BOLTON, G. E. and OCKENFELS, A. (2000), "Self-Centered Fairness in Games with More Than Two Players", *American Economic Review*, **90**, 166–193.

BRATBERG, E., TJOTTA, S. and OINES, T. (2005), "Do Voluntary International Environmental Agreements Work?", *Journal of Environmental Economics and Management*, **50**, 583–597.

BULOW, J., KLEMPERER, P. and GEANOKOPLOS, J. (1985), "Multimarket Oligopoly: Strategic Substitutes and Complements", *Quarterly Journal of Economics*, **93**, 488–511.

CARRARO, C. and SINISCALCO, D. (1993), "Strategies for the International Protection of the Environment", *Journal of Public Economics*, **52**, 309–328.

CESAR, H. S. J. and DE ZEEUW, A. (1996), "Issue Linkage in Global Environmental Problems", In A. Xepapadeas, (ed.), *Economic Policy for the Environment and Natural Resources: Techniques for the management and control of pollution*, (Cheltenham, UK: Edward Elgar) p. 158–173, ch. 7.

CONGLETON, R. D. (1992), "Political Institutions and Pollution Control", *Review of Economics and Statistics*, **74**, 412–21.

DE PAULA, A. (2009), "Inference in a Synchronization Game with Social Interactions", *Journal of Econometrics*, **148**, 56–71.

DE ZEEUW, A. (2008), "Dynamic Effects on the Stability of International Environmental Agreements", *Journal of Environmental Economics and Management*, **55**, 163–174.

DEMSETZ, H. (1967), "Toward a Theory of Property Rights", *American Economic Review*, **57**, 347–359.

EGGER and LARCH, M. (2008), "Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis", *Journal of International Economics*, **76**, 384–399.

EGGER, P. and MERLO, V. (2012), "Bits Bite: An Anatomy of the Impact of Bilateral Investment Treaties on Multinational Firms", *The Scandinavian Journal of Economics*, **114**, 1240–1266.

EINAV, L. (2010), "Not All Rivals Look Alike: Estimating an Equilibrium Model of the Release Date Timing Game", *Economic Inquiry*, **48**, 369–390.

ELKINS, Z., GUZMAN, A. T. and SIMMONS, B. A. (2006), "Competing for Capital: The Diffusion of Bilateral Investment Treaties 1960–2000", *International Organization*, **60**, 811–846.

EPA (1987), "Protection of Stratospheric Ozone", *Federal Register*, **52**, 47489, December 14.

FARRELL, J. and MASKIN, E. (1989), "Renegotiation in Repeated Games", *Games and Economic Behaviour*, **1**, 327–360.

FEENSTRA, R. C., LIPSEY, R. E., DENG, H., MA, A. C. and MO, H. (2005), "World Trade Flows: 1962–2000", (NBER Working Paper 11040, National Bureau of Economic Research, Cambridge, MA).

FEHR, E. and SCHMIDT, K. M. (1999), "A Theory of Fairness, Competition and Cooperation", *Quarterly Journal of Economics*, **114**, 817–868.

FERSHTMAN, C. and NITZAN, S. (1991), "Dynamic Voluntary Provision of Public Goods", *European Economic Review*, **35**, 1057–1067.

FINUS, M. and RUNDSHAGEN, B. (1998). "Renegotiation-proof Equilibria in a Global Emission Game When Players are Impatient", *Environmental and Resource Economics*, **12**, 275–306.

FINUS, M. and TJØTTA, S. (2003), "The Oslo Protocol on Sulfur Reduction: the Great Leap Forward?", *Journal of Public Economics*, **87**, 2031–2048.

FOLMER, H., v. MOUCHE, P. and RAGLAND, S. (1993), "Interconnected Games and International Environmental Problems", *Environmental and Resource Economics*, **3**, 313–335.

FREDRIKSSON, P. G. and GASTON, N. (2000), "Ratification of the 1992 Climate Change Convention: What Determines Legislative Delay?", *Public Choice*, **104**, 345–368.

FREUND, C. and ORNELAS, E. (2010), "Regional Trade Agreements", *Annual Review of Economics*, **2**, 139–166.

GEORGIADIS, G. (2015), "Projects and Team Dynamics", *Review of Economic Studies*, **82**, 187–218.

GOFFE, W., FERRIER, G. and ROGERS, J. (1994), "Global Optimization of Statistical Functions with Simulated Annealing", *Journal of Econometrics*, **60**, 69–99.

GROSSMAN, G. and HELPMAN, E. (1995), "The Politics of Free-trade Agreements", *American Economic Review*, **84**, 833–50.

GUZMAN, A. T. (1998), "Why LDCs Sign Treaties that Hurt Them: Explaining the Popularity of Bilateral Investment Treaties", *Virginia Journal of International Law*, **38**, 639–688.

HALLWARD-DRIEMEIER, M. (2003), "Do Bilateral Investment Treaties Attract Foreign Direct Investment? Only a Bit... and They Could Bite" (Policy Research Working Paper No. 3121, World Bank, Washington, DC).

HAMMITT, J. K. (2000), "Are the Costs of Proposed Environmental Regulations Overestimated? Evidence from the CFC Phaseout", *Environmental and Resource Economics*, **16**, 281–301.

HARSTAD, B. (2012), "Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations," *Review of Economic Studies*, **79**, 1527–1557.

HARSTAD, B. (2015). "The Dynamics of Climate Agreements", *Journal of the European Economic Association*. doi:10.1111/jeea.12138

HATHAWAY, O. A. (2002), "Do Human Rights Treaties Make a Difference?" *Yale Law Journal*, **111**, 1935–2042.

HEAD, K., MAYER, T. and RIES, J. (2010), "The Erosion of Colonial Trade Linkages after Independence", *Journal of International Economics*, **81**, 1–14.

HOEL, M. and SCHNEIDER, K. (1997), "Incentives to Participate in an International Environmental Agreement", *Environmental and Resource Economics*, **9**, 153–70.

HOEL, M. (1992), "International Environmental Conventions: The Case of Uniform Reductions of Emissions", *Environmental and Resource Economics*, **2**, 141–159.

HONORÉ, B. E. and DE PAULA, A. (2010), "Interdependent Durations", *Review of Economic Studies*, **77**, 1138–1163.

IMF (2015), "Review of the Role of Trade in the Work of the Fund", (Policy paper, International Monetary Fund, Washington, DC).

JAIMOVICH, D. (2012), "A Bayesian Spatial Probit Estimation of Free Trade Agreement Contagion", *Applied Economics Letters*, **19**, 579–583.

KERNER, A. (2009), "Why Should I Believe You? the Costs and Consequences of Bilateral Investment Treaties", *International Studies Quarterly*, **53**, 73–102.

KESSING, S. G. (2007), "Strategic Complementarity in the Dynamic Private Provision of a Discrete Public Good", *Journal of Public Economic Theory*, **9**, 699–710.

LANGE, A. and VOGT, C. (2003), "Cooperation in International Environmental Negotiations Due to a Preference for Equity", *Journal of Public Economics*, **87**, 2049–2067.

LEE, B. and INGRAM, B. F. (1991), "Simulation Estimation of Time-series Models", *Journal of Econometrics*, **47**, 197–205.

MAYER, T. and ZIGNAGO, S. (2011), "Notes on CEPII's Distances Measures: The GeoDist Database", (Working Paper 2011–25, CEPII, Paris, France).

McFADDEN, D. (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration", *Econometrica*, **57**, 995–1026.

McGINTY, M. (2007), "International Environmental Agreements Among Asymmetric Nations", *Oxford Economic Papers*, **59**, 45–62.

MILGROM, P. and ROBERTS, J. (1990), "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities", *Econometrica*, **58**, 1255–1277.

MINA, W. (2013), "Beyond FDI: The Influence of Bilateral Investment Treaties on Debt", (International Center for Public Policy Working Paper 13–25, Georgia State University, Atlanta, GA).

MURDOCH, J. C. and SANDLER, T. (1997), "The Voluntary Provision of a Pure Public Good: The Case of Reduced CFC Emissions and the Montreal Protocol", *Journal of Public Economics*, **63**, 331–349.

MURDOCH, J. C., SANDLER, T. and VIJVERBERG, W. P. (2003), "The Participation Decision Versus the Level of Participation in an Environmental Treaty: A Spatial Probit Analysis", *Journal of Public Economics*, **87**(2), 337–362.

NBER-United Nations Trade Data (1962-2000), "World Trade Flows" http://cid.econ.ucdavis.edu/nberus.html (accessed May 2012).

NEUMAYER, E. and PLÜMPER, T. (2010), "Spatial Effects in Dyadic Data", *International Organization*, **64**, 145–166.

NEUMAYER, E. (2002), "Does Trade Openness Promote Multilateral Environmental Cooperation?", *World Economy*, **25**, 815–832.

PAKES, A. and POLLARD, D. (1989), "Simulation and the Asymptotics of Optimization Estimators", *Econometrica*, **57**, 1027–1057.

PANAGARIYA, A. (1999), "The Regionalism Debate: An Overview", *World Economy*, **22**, 455–476.

PARSON, E. A. (2003), *Protecting the Ozone Layer. Science and Strategy* (Oxford, UK: Oxford University Press).

PEARCE, D. G. (1990), "Repeated Games: Cooperation and Rationality", In J.-J. Laffont, (ed.), *Advances in Economic Theory, Sixth World Congress*, Vol. 1, (Cambridge, UK: Cambridge University Press), 132–174, ch. 4.

RUBIO, S. J. and ULPH, A. (2007), "An Infinite-Horizon Model of Dynamic Membership of International Environmental Agreements", *Journal of Environmental Economics and Management*, **54**, 296–310.

SAMUELSON, P. A. (1954), "The Pure Theory of Public Expenditure", *The Review of Economics and Statistics*, **36**, 387–389.

SCHMIDT-DENGLER, P. (2006), "The Timing of New Technology Adoption: The Case of MRI", (Working paper, London School of Economics, London, UK).

SEBENIUS, J. K. (1983), "Negotiation Arithmetic: Adding and Subtracting Issues and Parties", *International Organization*, **37**, 281–316.

SUNSTEIN, C. (2007), "Of Montreal and Kyoto: A Tale of Two Protocols", *Harvard Environmental Law Review*, **31**, 1–66.

SWEETING, A. (2006), "Coordination, Differentiation and the Timing of Radio Commercials", *Journal of Economics and Management Strategy*, **15**, 909–942.

TOBIN, J. and ROSE-ACKERMAN, S. (2003), "Foreign Direct Investment and the Business Environment in Developing Countries: The Impact of Bilateral Investment Treaties", (William Davidson Institute Working Paper 587, University of Michigan Business School, Ann Arbor, MI).

TOBIN, J. L. and BUSCH, M. L. (2010), "A Bit is Better Than a Lot: Bilateral Investment Treaties and Preferential Trade Agreements", *World Politics*, **62**, 1–42.

TOLLISON, R. D. and WILLETT, T. D. (1979), "An Economic Theory of Mutually Advantageous Issue Linkage in International Negotiations", *International Organization*, **33**, 425–449.

TOPKIS, D. (1979), "Equilibrium Points in Nonzero-Sum *n*-Person Submodular Games", *SIAM Journal of Control and Optimum*, **17**, 773–787.

TOPKIS, D. M. (1998), *Supermodularity and Complementarity*. Frontiers in Economic Analysis. (Princeton, NJ: Princeton University Press).

UNCTAD (1959-2015). "International Investment Agreements Navigator" http://investmentpolicyhub.unctad.org/iia, United Nations (accessed on April 19, 2015).

UNEP (1986-2015), “GEO Data Portal” http://geodata.grid.unep.ch/, United Nations Environment Programme (accessed on March 27, 2007).

UNEP (1999), *Handbook on Data Reporting under the Montreal Protocol*. (Paris, France: United Nations Environment Programme).

VINER, J. (1950), *The Customs Union Issue*. (New York: Carnegie Endowment for International Peace).

VIVES, X. (1990), “Nash Equilibrium with Strategic Complementarities”, *Journal of Mathematical Economics*, **19**, 305–321.

WAGNER, U. J. (2001), “The Design of Stable International Environmental Agreements: Economic Theory and Political Economy”, *Journal of Economic Surveys*, **15**, 377–411.

WAGNER, U. J. (2009), “The Voluntary Provision of a Pure Public Good? Another Look at CFC Emissions and the Montreal Protocol”, *Oxford Economic Papers*, **61**, 183–196.