Estimating Strategic Complementarities in a Dynamic Game of Timing: The Case of the Montreal Protocol

Ulrich J. Wagner*

JOB MARKET PAPER

January 14, 2008

Abstract

The Montreal Protocol on Substances that Deplete the Ozone Layer is widely regarded as one of the greatest successes of international environmental policy. By controlling the stock of stratospheric pollutants that ultimately increase the prevalence of skin cancer, this treaty has accomplished the provision of a pure public good at the global scale. This paper estimates the magnitude of social reinforcement effects that reined in free-riding by individual countries and rendered ratification of the treaty a strategic complement. Since a non-strategic analysis of the ratification process fails to account for forward-looking behavior by governments, I develop a strategic model of the timing of treaty ratification which is amenable to structural estimation. The model predicts that strategic complementarities accelerate ratification, to the extent that ratification by one country may trigger ratification by another. I exploit this prediction to identify strategic complementarities in a structural econometric model of ratification of the Montreal Protocol. The fitted model predicts that strategic complementarities reduced the average time to ratification by 35 weeks (one fifth of the standard deviation). Investigation of different sources of strategic complementarities points to concern about reputation and to a preference for equity. There is no evidence that bilateral trade flows affected the strength of strategic complementarities.

*The Earth Institute, Columbia University. Web: http://www.columbia.edu/∼uw2101. This paper is based on the first two chapters of my Ph.D. dissertation in economics at Yale University. I would like to thank my committee members Pat Bayer, Steve Berry, and Chris Timmins for advice and support. I have benefitted from discussions with Scott Barrett, Dirk Bergemann, Hanming Fang, Geoff Heal, Nat Keohane, Gerard Llobet, Pei-Yu Lo, Pedro Mira, Bernard Salanié and Philipp Schmidt-Dengler. Participants at the Yale environmental economics seminar and the microeconomics prospectus workshop as well as attendants of the 2007 North American Summer Meeting of the Econometric Society have provided many helpful comments. All errors are my own. 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 Leitner Grant in International and Comparative Political Economy from Yale University.

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 fields of international politics, including military defense, 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.

Failure of the international community to provide global public goods or to protect global common property resources is often attributed to the adverse incentives faced by individual governments. Clearly, unless every country takes into account the external benefits (i.e. the benefits accruing to all other countries) of its 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. Similarly, unregulated access to a common property resource provides an incentive for over-exploitation since a country cannot expect to recoup the cost of investments into conservation (Hardin, 1968).

At the national level, such conflicts between individual and collective rationality can be resolved by the intervention of the government (Demsetz, 1967), as evidenced by the variety of public goods provided in most developed countries. 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[1] 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. In so doing, sovereign states seek to coordinate their policies in mutually beneficial ways. By the very nature of sovereignty, however, such agreements are fundamentally non-binding, as states can always withdraw from an agreement.

This is why most of the theoretical economic literature uses non-cooperative game theory to analyze the incentives for participation in international agreements (e.g. Barrett, 1994).

[1]The legal procedure of a state joining a multilateral agreement is the signature followed by ratification, which marks the legal accession. In this paper, 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.

Governments are treated as monolithic decision makers who accede to and comply with a treaty only if doing so is to their national benefit. In contrast, empirical studies of international treaties pay little attention to the role of strategic interaction (e.g. Congleton, 1992), with a few notable exceptions (Auffhammer et al., 2005). This paper bridges the gap by developing a strategic model of treaty formation which is amenable to structural estimation. The model is firmly rooted in the non-cooperative paradigm but gives particular attention to the role of social reinforcement effects in enhancing cooperation. The framework is applied to estimate empirically the magnitude of such effects using data on the ratification of the Montreal Protocol on Substances that Deplete the Ozone Layer.

The Montreal Protocol was opened for signature in 1987 and has since been ratified by 191 countries. By controlling the stock of stratospheric pollutants that ultimately increase the prevalence of skin cancer, this international environmental agreement has overcome countries' incentive to free-ride and accomplished the provision of a pure public good at the global scale. The treaty is widely considered to be one of the greatest successes in the history of international environmental policy. This paper argues that cooperation under the Montreal Protocol is enhanced by government interaction effects that render ratification a strategic complement (Bulow et al., 1985), in the sense that ratification by one country increases the relative benefits to ratification for others. Strategic complementarity arises, for example, if a government (or its political constituency) has a preference for equity and the Montreal Protocol is perceived as more equitable the more other countries – especially large polluters – contribute their "fair" share of abatement. Similarly, the government of a non-member state may suffer a reputation loss that is more severe if participation in the treaty is broad (Hoel and Schneider, 1997). Another possible source of strategic complementarity is the treaty's ban of trade in controlled substances between signatories and non-signatories. This is because the potential gains from trade reinforce the incentive for participation once membership in the treaty is sufficiently large (Barrett, 1997b).

Empirical analysis of strategic interaction in international treaty-making is complicated by the fact that the ratification process is not repeatedly observed. This paper develops a novel econometric framework that exploits the variation in ratification dates to estimate the magnitude of strategic complementarities. To this end, I first present a general theoretical model of the timing of public good provision. I analyze a voluntary provision game among asymmetric players which is repeated infinitely often. In each period, players choose whether to cooperate or to defect. The relative payoff to cooperation – that is, the difference in payoffs to both actions – is assumed to increase monotonically over time, reflecting changes in both the provision cost and the valuation of the public good. Since ozone depletion is caused by stock pollutants that accumulate with a delay of several decades, the evolution of abatement benefits during the ratification period is regarded as predetermined by the trend in historical emissions. Strategic complementarity in public good provision means that the relative payoff to cooperation increases with the number of players who cooperate. I solve for the unique strongly renegotiation-proof Nash equilibrium of the game and show that strategic complementarity brings forward public good provision in time, although provision remains slower than socially optimal. If time periods are sufficiently small, the model predicts that clustering of provision decisions in time can only occur in the presence of strategic complementarities.

This prediction helps to identify strategic complementarities in a structural econometric model of ratification of the Montreal Protocol. While the ratification dates of individual countries trace out the distribution of the private net cost, the frequency of multiple countries ratifying on the same day reveals the strength of strategic complementarities. Using the method of simulated moments, I estimate both these components of the payoff function and find evidence of strategic complementarities in treaty ratification. A counterfactual experiment shows that strategic complementarities reduced the average time to ratification by 35 weeks (about one fifth of the standard deviation).

To further explore this phenomenon, I investigate specific reinforcement effects at the intergovernmental level, which have been suggested in the economics and international relations literatures. I find that ratification by country A is more likely to trigger ratification by country B if both countries have ratified pre-existing international treaties, or if country A has a large share in global emissions of ozone-depleting substances. I interpret the former as indicating concern about reputation and the latter as reflecting a preference for equity. By contrast, I do not find evidence that bilateral trade flows affected the strength of strategic complementarities.

The remainder of the paper is structured as follows. The next section reviews the economic literatures on timing games on the one hand and on international environmental agreements on the other hand. Section 3 develops a general theoretical model framework of the timing of public good provision and derives implications for structural estimation. Section 4 provides background information on the science and diplomacy underlying the Montreal Protocol. Section 5 presents a structural econometric model of the ratification of the Montreal Protocol. Section 6 describes the data set and performs the estimation of the baseline model. The role of different sources of strategic complementarities is investigated in section 7. Section 8 concludes.

2 Literature Review

2.1 Timing Games

The model presented in this paper ties together the literature on non-cooperative games of timing on the one hand and the literature on supermodular games and strategic complementarities on the other hand. I draw on insights from both strands to shed light on the role of endogenous effects in the timing of public good provision within a non-cooperative environment. Furthermore, the model is related to the literature on social interactions in that it focuses on interactions between a large number of players.

Game theorists have studied a great variety of non-cooperative games of timing. Among the better-known timing games are the war of attrition (Smith, 1974), patent races (Fudenberg et al., 1983), pre-emption games (Reinganum, 1981a,b, Fudenberg and Tirole, 1985) and exit of oligopolist firms in declining industries (Ghemawat and Nalebuff, 1985). Hendricks et al. (1988) provide a unified treatment of these games, all of which assume that the effect of a player's action on other players' payoffs is negative. This is in contrast to the model presented below where players' actions are strategic complements.

Strategic complementarity in the payoff function is at the heart of the theory of supermodular games, introduced by Topkis (1979, 1998) and further developed in Vives (1990), Milgrom and Roberts (1990), and Zhou (1994). This theory establishes results regarding existence, general properties and comparative statics of pure-strategy Nash equilibria in supermodular games on which I will draw in the subsequent analysis. A fundamental result is that strategic complementarities in a game of complete information may lead to multiple Nash equlibria that are Pareto-ranked (Cooper and John, 1988). An analogous finding derived in the literature on non-market interactions is that social multipliers may induce large discrepancies in outcomes in spite of small differences in economic fundamentals (see Glaeser and Scheinkman, 2002, and the literature cited therein). Brock and Durlauf (2001) analyze such interactions in the context of a discrete-choice model with incomplete information and discuss aspects relating to the identification and estimation of this type of model.

Recently, timing games have received more attention in empirical work, mainly in the field of industrial organization.[2] Research in this area can be categorized along several dimensions.

[2]Einav (2003) studies a Bayesian game of the timing of release dates for new movies and finds that motion picture distributors cluster release dates too much compared to dates that would maximize joint profits. Schmidt-Dengler (2006) estimates a dynamic model of the adoption of MRI technology in US hospitals that focuses on two different sources of strategic interaction: preemption and business-stealing motives. As in Einav (2003), the business stealing effect gives players an incentive to differentiate their actions over time. Sweeting (2006) develops an estimable game theoretical model of the timing of radio commercials that allows both coordination and differentiation over time. In another paper, Sweeting (2005) studies the incentives

First, a distinction can be made between static (Einav, 2003, Sweeting, 2005, 2006) and dynamic games of timing (Schmidt-Dengler, 2006, De Paula, 2006). Second, the nature of the externality is a fundamental aspect of the strategic situation that is analyzed. Like the theoretical papers cited above, most empirical studies focus on environments with negative externalities such as oligopoly settings (Einav, 2003, Sweeting, 2006, Schmidt-Dengler, 2006). Third, the information structure varies across studies. In some economic settings, private information about individual payoffs is relevant. Hence equilibrium strategies depend on the expected actions taken by other players (Einav, 2003, Sweeting, 2005, Brock and Durlauf, 2001, De Paula, 2006).

Not least, different model frameworks place different requirements on the data for the purposes of estimation. For example, there is a closely related literature on estimating dynamic models of entry and exit, pioneered by Aguirregabiria and Mira (2007), Pakes et al. (2007), Pesendorfer and Schmidt-Dengler (2006), Bajari et al. (2007), and applied by Ryan (2006) and Collard-Wexler (2006). These estimators require the data set to be sufficiently rich so as to allow the researcher to estimate state transition probabilities and policy functions for all players. The timing game studied in Schmidt-Dengler (2006) is less demanding on the data because the decision to adopt is irreversible over time.

I extend this literature by analyzing a discrete dynamic game of complete information where positive externalities render actions strategic complements. Similar to De Paula (2006), who also studies a game with strategic complementarities, I find that clustering occurs as a consequence of endogenous effects. His analysis and mine are complementary in the sense that they derive this result starting with fundamentally different assumptions about the information structure (incomplete vs. complete), the time structure (continuous vs. discrete), the strategy space (irreversible vs. reversible actions), and the solution concept. My framework provides researchers who are looking to go beyond a static model with an alternative that is feasible in settings with limited data availability, including the case where only a single history of the game is observed.

facing stations to coordinate the timing of commercials in a game with private information. He tests for the presence of multiple Bayesian Nash equilibria and recovers estimates of the structural parameters of the game using all of the equilibrium outcomes observed in the data. He finds moderate incentives for coordination. In independent work, De Paula (2006) studies a continuous-time "synchronization game" with social interactions. Players only observe their own payoffs and whether or not other players have dropped out. De Paula shows that simultaneous stopping by multiple agents occurs if and only if payoffs exhibit strategic complementarity, provided that the stochastic evolution of payoffs over time is not subject to negative jumps. He uses this feature to devise a test for the presence of endogenous effects with discrete data. He applies this methodology to analyze desertion during the American Civil War.

2.2 Economic Analysis of International Environmental Agreements

The number of international environmental agreements (IEA) in the world that are currently in force or expected to enter into force in the short term is estimated at close to 300 (Barrett, 2003). The majority of them is aimed at protecting a shared common resource or providing a public good at the regional or global scale. The incentives that countries face to participate in such an agreement and the question of how effective it can be, given the constraint that no country can be coerced into participating, are studied in a large body of theoretical work.[3] The modeling approaches taken in this literature differ widely, ranging from the standard non-cooperative game of public good provision due to Bergstrom et al. (1986) to sophisticated burden-sharing rules inspired by cooperative game theory (Chander and Tulkens, 1995, and others). The most popular approach casts participation in a so-called *self-enforcing international environmental agreement* as a non-cooperative game of coalition formation (see Barrett, 1994, Carraro and Siniscalco, 1993, Hoel, 1992, Kolstad, 2007, Rubio and Ulph, 2006, and others).

A self-enforcing IEA is modeled as a static game in three stages. Countries first decide whether to accede or not. Next, signatory countries determine their abatement levels so as to maximize joint payoffs. Finally, non-parties unilaterally choose their optimal abatement levels, taking abatement by all other countries as given. In its most basic form, the model shows that there is a trade-off between maximizing participation and the environmental effectiveness of a self-enforcing IEA. That is, individual incentives to defect are strongest when aggregate benefits to full cooperation are high, resulting in low participation rates. The introduction of additional model features can change this result. For the Montreal Protocol, Barrett (1997b) shows that the ban of trade in controlled substances between member states and non-member states transforms accession to a self-enforcing IEA into a coordination game.

While the theoretical literature on IEAs has been growing rapidly over the past 15 years or so, there are only few empirical studies of IEAs. With respect to the Montreal Protocol, Congleton (1992) presents econometric evidence that democracies were more likely to sign the Montreal Protocol than autocracies. Murdoch and Sandler (1997) regress reductions in chlorofluorocarbon (CFC) emissions in the run-up to the treaty on GNP using least-squares techniques. They interpret their finding of positive and linear relationship as evidence in support of a non-cooperative model of emission reductions. Upon re-examination of the evidence presented in that paper, I conclude that the strong correlation between GNP and emission reductions is most likely induced by the procedure that was used to impute between

[3] I provide a comprehensive survey of this literature elsewhere (Wagner, 2001).

80% and 90% of the CFC emission data (Wagner, 2007). In a replication of the econometric analayis using self-reported emission data from UNEP, I must reject the hypothesis of a positive and linear relationship between income and abatement.

Beron et al. (2003) study interdependencies in the decision to ratify the Montreal Protocol. They estimate a Probit model where the latent benefits of joining the treaty depend on the latent benefits of all other countries. Identification of this spillover effect comes from data on bilateral trade flows or individual shares in world CFC consumption. The estimated spillover coefficients are not statistically significant. Auffhammer et al. (2005) put more structure on the nature of strategic interaction between countries by modeling CFC production as a Cournot duopoly between US producers and the rest of the world. Using time series data on CFC production, they examine whether any of the two players tried to increase emissions in order to improve their bargaining position in the run-up to the Montreal negotiations. They find evidence of strategic behavior by producers outside the United States.[4]

The analysis in this paper contributes to both the theoretical and the empirical literatures in several ways. First, it presents the first strategic model of the *timing* of ratification. The model explains why ratification decisions by some countries are spaced out over time and others are clustered on the same day. Moreover, the model easily accommodates heterogenous payoff functions with asymmetric interaction terms. In the self-enforcing IEA model, payoff heterogeneity is difficult to deal with, leading to multiple equilibria and often preventing an analytical solution (Barrett, 1997a, McGinty, 2007).

An important advantage of my model is that it is amenable to structural estimation. The main problem with the empirical implementation of the self-enforcing IEA is that a given agreement corresponds to a single observation, namely the participation rate in that agreement. Clearly, such a model is not estimable. Therefore, empirical studies have focused on reduced-form analysis (Murdoch and Sandler, 1997, Beron et al., 2003), limited the number of agents to two stylized players (Auffhammer et al., 2005) or ignored strategic interaction altogether (Congleton, 1992, Hathaway, 2004).

By developing an estimable strategic model of timing, I attempt to provide a unified framework for empirical analysis that bridges the gap between the theoretical and empirical literatures. The model focuses on interdependent ratification decisions and extends the work by Beron et al. (2003). First, the dynamic model setup with complete information allows players to condition their actions on observed participation rather than on unobserved latent

[4]Econometric work on other international environmental treaties has focused on the Convention on Long-Range Transboundary Air Pollution. Murdoch et al. (1997) use a similar approach as Murdoch and Sandler (1997) to study reductions of sulfur dioxide and nitrogen oxide emissions. Murdoch et al. (2003) estimate a two-stage model of participation and abatement decisions in the Helsinki Protocol that regulates emissions of sulfur.

benefits as in Beron et al. (2003). Furthermore, since endogenous effects are identified off the time intervals between subsequent ratification decisions, I do not need to assume a specific political or economic channel through which interactions work – though I can and will test whether some such channels are relevant. Finally, estimation of a structural rather than a reduced-form model of strategic interaction allows me to perform counterfactual experiments.

3 A General Model of the Timing of Public Good Provision

This section introduces a dynamic model of the timing of public good provision. The basic setup is a voluntary contribution game which is repeated infinitely often while the relative benefit to cooperation gradually increases over time. As I argue below, this model captures the basic strategic choices facing countries in the ratification of the Montreal Protocol. I characterize equilibrium in this model with and without strategic complementarities. Next, I discuss the efficiency properties of non-cooperative equilibrium and examine its limit as the length of time period goes to zero. The section concludes with a discussion of the model's implications for empirical analysis.

3.1 Model Setup

3.1.1 Preliminaries

Let $\Gamma(t) = (I, A, \pi(t))$ be an N-player game, where $I = \{1, \dots, N\}$ is the set of players, $A_i = \{0, 1\}$ is the action set of player $i \in I$, $A = \prod_{i \in I} A_i$ denotes the action space, $t \in T = \{0, 1, \dots\}$ is a parameter, $\pi_i(t) : A \times T \mapsto \mathbb{R}$ is the payoff function for player i, and $\pi \equiv (\pi_1, \pi_2, \dots, \pi_N)$ is the payoff function for the game. The generic element $a = (a_i, a_{-i})$ of A is an action profile that consists of action $a_i \in A_i$ taken by player i and the vector of actions taken by all players other than i, $a_{-i} \in \prod_{j \neq i} A_j \equiv A_{-i}$. Actions 1 and 0 will be referred to as "cooperate" and "defect", respectively.

Next, consider the infinite-horizon game $G = \{I, S, V(0)\}$ formed by the sequence of stage games $\{\Gamma(t)\}_{t \in T}$, where I is the set of players, $S_i = \prod_{t \in T} A_i$ is the strategy set of player $i \in I$, $S = \prod_{i \in I} S_i$ is the set of feasible strategy profiles, $V_i(0) : S \to \mathbb{R}$ is the payoff function for player i, and $V(0) \equiv (V_1(0), V_2(0), \dots, V_N(0))$ is the payoff function for the game. Define a history $h(t)$ of the game $G(0)$ to be a sequence of action profiles $\{a(\tau)\}_{\tau=0}^{t-1}$ where $a(\tau) \in A$ is the action profile played at node τ. The collection of all possible histories $H(t)$ at node t is given by $\prod_{\tau=0}^{t-1} A$. Define $H = \bigcup_{t=0}^{\infty} H(t)$ as the collection of all possible histories. A strategy $s_i : H \mapsto A_i$ is a function that designates a particular action for player i for any given history $h \in H$. Define the continuation game $G(t) = (I, S_t, V(t))$ as the information set $h(t)$ and all subsequent decision nodes of the game $G(0)$. Denote by $S_i^t = \prod_{\tau=t}^{\infty} A_i$ the strategy set of player i and by $S_t = \prod_{i \in I} S_i^t$ the set of strategy profiles in the continuation game $G(t)$. Player i's continuation payoff in period t under strategy profile s is given by

$$V_i(s,t) = \sum_{\tau=t}^{\infty} \pi_i(s_i(\tau), s_{-i}(\tau), \tau) e^{-r_i(\tau - t)} \tag{3.1}$$

where $s(t) = (s_i(t), s_{-i}(t))$ is the action profile induced at time t by strategy profile s and r_i is player i's discount rate. The elements of the payoff function $V_i(0)$ are given by evaluating eq. (3.1) at $t = 0$ for each $i \in I$.

3.1.2 The Payoff Function

To further characterize the primitives of the model, the following assumptions are made on the per-period payoff function of the stage game Γ. First, the *relative payoff to cooperation*, defined as

$$\Delta\pi_i(a_{-i}, t) \equiv \pi_i(1, a_{-i}, t) - \pi_i(0, a_{-i}, t) \tag{3.2}$$

is required to be a strictly increasing, twice continuously differentiable function of calendar time $\tilde{t} \in \mathbb{R}_+$.

Assumption 1 (Strictly increasing differences in calendar time $\tilde{t}$) *$\forall i \in I$, $\forall \tilde{t} \in \mathbb{R}_+$, $\forall a_{-i} \in A_{-i}$;*

$$\Delta\pi_i(a_{-i}, \tilde{t}) \in \mathrm{C}^2 \quad \textit{and} \quad \frac{\partial \Delta\pi_i(a_{-i}, \tilde{t})}{\partial \tilde{t}} > 0$$

Assumption 2 (Positive spillovers) *$\forall i \in I$, $\forall t \in T$, $\forall a \in A$ such that $a_{-i} \leq a'_{-i}$, $a_{-i} \neq a'_{-i}$;*

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

Assumption 2 ensures that cooperation by any one player creates a strictly positive external benefit for the other players. In addition, I consider two alternative ways in which cooperation by other players affects the relative payoffs to cooperation:

Assumption 3 (Constant differences) *$\forall i \in I$, $\forall t \in T$, $\forall a_{-i}, a'_{-i} \in A_{-i}$ such that $a_{-i} \leq a'_{-i}$, $a_{-i} \neq a'_{-i}$;*

$$\pi_i(0, a'_{-i}, t) - \pi_i(0, a_{-i}, t) = \pi_i(1, a'_{-i}, t) - \pi_i(1, a_{-i}, t)$$

Assumption 4 (Increasing differences) *$\forall i \in I$, $\forall t \in T$, $\forall a_{-i}, a'_{-i} \in A_{-i}$ such that $a_{-i} \leq a'_{-i}$, $a_{-i} \neq a'_{-i}$;*

$$\pi_i(0, a'_{-i}, t) - \pi_i(0, a_{-i}, t) \leq \pi_i(1, a'_{-i}, t) - \pi_i(1, a_{-i}, t)$$

and the inequality is strict for some $j \in I$.

Under assumption 3, the difference in payoffs to both actions is invariant with respect to the number of other players who cooperate. This is the case, for example, if the marginal benefit of the public good is constant. By contrast, assumption 4 allows for the possibility that the relative payoff to cooperating may increase with the number of other players who cooperate. This effect is henceforth referred to as a *strategic complementarity.*

Assumption 5 (Timing game) $\pi_i(1, \mathbf{0}, 0) < \pi_i(0, \mathbf{0}, 0) \quad \forall i \in I.$

Assumption 6 (Boundedness) $\lim_{t\to\infty} e^{-r_i t}\pi_i(a, t) = 0 \quad \forall i \in I.$

Assumption 5 precludes degenerate timing games in which one or more players have a dominant strategy to cooperate unilaterally from the beginning of the game and assumption 6 ensures that the game G has a well-defined solution.

3.1.3 Equilibrium

The solution concepts commonly used in the analysis of static and dynamic games of complete information are Nash equilibrium and subgame-perfect Nash equilibrium, respectively. In a Nash equilibrium, no player can benefit by changing her action given the actions taken by all other players:

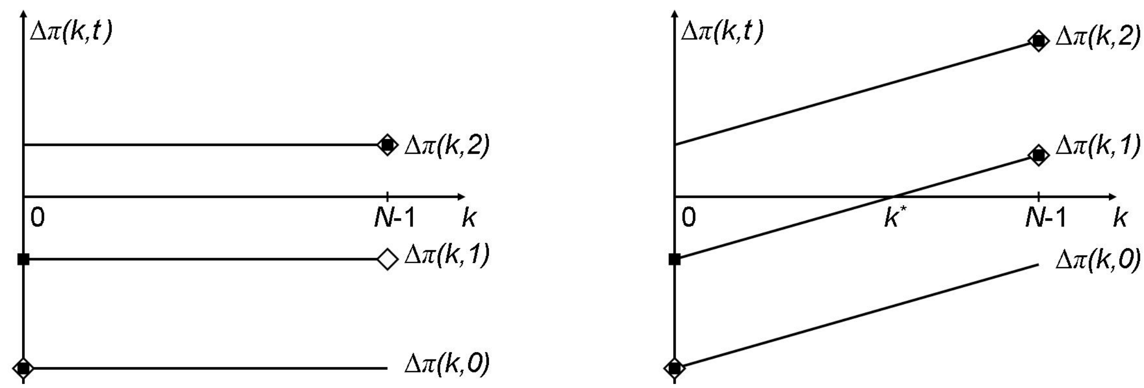
Definition 1 (Nash equilibrium) *An action profile $\tilde{a} \in A$ is a Nash equilibrium (NE) of the stage game $\Gamma(t)$ if $\forall i \in I, \forall a_i \in A_i$,*

$$\pi_i(\tilde{a}_i, \tilde{a}_{-i}, t) \geq \pi_i(a_i, \tilde{a}_{-i}, t).$$

In a subgame-perfect Nash equilibrium of the infinite-horizon game, no player can obtain a higher continuation payoff by changing her strategy, given the strategies adopted by all other players.

Definition 2 (Subgame-perfect Nash equilibrium) *A profile $\tilde{s} \in S_t$ is a subgame-perfect Nash equilibrium (SPNE) of the game $G(t)$ if $\forall i \in I$, $\forall s_i \in S_i^t, \forall \tau \geq t$*

$$V_i(\tilde{s}_i, \tilde{s}_{-i}, \tau) \geq V_i(s_i, \tilde{s}_{-i}, \tau).$$



(a) Constant differences (b) Increasing differences

Figure 1: Stage game payoffs with identical players

To obtain sharper predictions on the equilibrium play, I shall focus on subgame-perfect Nash equilibria that are renegotiation-proof. The precise definition of this refinement concept is given in the next section.

3.2 The Timing of Voluntary Provision

3.2.1 The case of symmetric players

To provide some intuition for the model, I first sketch the solution for the case of identical players. The relative payoff to cooperation – defined in eq. (3.2) as the difference between the payoffs to cooperation and defection – can hence be written in terms of the number of other players k who cooperate in any given period t. Figure 1 depicts how this function $\Delta\pi(k,t)$ changes with its arguments assuming either constant or increasing differences in the payoff functions.

Consider first the case of constant differences. Figure 1a depicts the function $\Delta\pi(k,t)$ as a flat line for each of the stage games played in the first three periods 0, 1 and 2. This line shifts upward as time passes because the relative payoff to cooperation increases over time (assumption 1). In periods 0 and 1, the dominant strategy for every player is to defect because cooperation gives a strictly lower payoff regardless of other players' actions. Conversely, in period 2 every player chooses to cooperate irrespective of the number of other players who cooperate. Therefore, each of the stage games depicted here has a unique Nash equilibrium in dominant strategies. This outcome is marked by the symbol ■. Furthermore, the cooperative Nash equilibrium in period 2 is efficient because it maximizes the external benefits from public good provision. The efficient outcome is marked by the symbol $\diamond$. If the total costs of public good provision outweigh the total benefits, the non-cooperative Nash equilibrium is efficient. This is assumed in period 0. However, as the net costs of cooperation fall over time, the non-cooperative Nash equilibrium may result in underprovision of the public good, as depicted in period 1. This is the familiar prisoners' dilemma where external benefits of public good provision are not internalized by individually rational players.

Figure 1b depicts the stage game when payoffs exhibit increasing differences. In contrast to figure 1a, the relative payoff to cooperation is now increasing in k, the number of other players who cooperate. This reinforcement effect creates an additional Nash equilibrium in period 1: In addition to the non-cooperative outcome, full cooperation is a Nash equilibrium, too. This is because each player i is now strictly better off by cooperating if at least k^* other players cooperate. Conversely, given a positive number of contributors lower than k^*, each contributor can secure a strictly higher payoff by defecting. Cooperation is a strategic complement in this game, and a critical mass of k^* players can tip the equilibrium from a situation with no public good provision to one with efficient provision.[5] The dominant strategy equilibria in periods 0 and 2 remain unchanged.

Consider now the multi-stage game comprised of periods 0, 1, and 2. Using backward induction, we know that a subgame-perfect Nash equilibrium of the multi-stage game requires that a Nash equilibrium be played at every stage. Therefore, the unique subgame-perfect Nash equilibrium for the case of constant differences is given by the sequence (defect, defect, cooperate) for each player. If payoffs are increasing in the provision decisions of other players, there is an additional subgame-perfect Nash equilibrium induced by the profile (defect, cooperate, cooperate). Since defection in period 1 foregoes the external benefits of public good provision, the latter equilibrium Pareto-dominates the former.

Multiple equilibria are a well-known property of supermodular games – that is, games with strategic complementarities (see Cooper and John, 1988, Brock and Durlauf, 2001, Glaeser and Scheinkman, 2002). In the infinite-horizon game, multiplicity of equilibrium is further exacerbated because *any* individually rational payoff vector can be sustained in subgame-perfect Nash equilibrium (Fudenberg and Maskin, 1986). Multiple equilibria pose a problem in that they limit the predictive power of the game theoretical model. In an empirical framework, this complicates econometric estimation of the model parameters because there is no one-to-one mapping from the data into the parameter space.[6]

[5] See Heal and Kunreuther (2007) for a formal analysis of tipping of equilibria in normal-form games with increasing differences. These authors also provide a rich set of examples of such games in economics and sociology.

[6] To circumvent this problem, researchers often impose additional assumptions that result in sharper predictions on equilibrium play. For example, many authors assume that players move in a particular order

Multiplicity of equilibrium vanishes in this game if one requires that equilibrium outcomes not be vulnerable to renegotiation. While in a (subgame-perfect) Nash equilibrium no player stands to gain from a unilateral deviation, there may still be mutually beneficial deviations for groups of players. For example, in figure 1b players have an incentive to jointly renege on the non-cooperative NE and move towards the cooperative NE in period 1. Given a setting with unconstrained communication among $N > 2$ players, it seems plausible to require the solution to be robust to such collective deviations. This is the idea underlying the concept of *coalition-proof Nash equilibrium* (CPNE), a refinement of Nash equilibrium for strategic form games (Bernheim et al., 1987). An appealing aspect of CPNE is that a strategy profile need not be robust to deviations by subsets of players unless those deviations are *self-enforcing*.[7] What is more, CPNE is Pareto-efficient in the class of self-enforcing profiles. Hence, CPNE eliminates the non-cooperative outcome in the stage game depicted in figure 1b. Unfortunately, this concept is not defined for infinite-horizon games.[8] The closest analogue to CPNE is the concept of renegotiation-proofness which focuses on subgame-perfect strategy profiles that are robust to joint deviations by *all* (but not: sub-coalitions of) players. Several alternative definitions of renegotiation-proofness have been suggested.[9] This paper follows the literature on the stability of multilateral treaties (see e.g. Barrett, 1994, 2002, Finus, 2001, and the literature cited therein) in employing the concepts of weak and strong renegotiation-proofness due to Farrell and Maskin (1989a). The definitions are as follows.

Definition 3 (Weakly Renegotiation-Proof Equilibrium) *(Farrell and Maskin, 1989a) 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. If an equilibrium s is WRP, then we also say that the payoffs $V(s)$ are WRP.*

that gives rise to a unique equilibrium (see, e.g. Berry, 1992, Mazzeo, 2002, Schmidt-Dengler, 2006). Others pick a Nash equilibrium with certain properties, for example, one that maximizes joint profits of all players (Sweeting, 2006) or one that is robust to small perturbations in the information structure (De Paula, 2006). Jia (2005) estimates an entry game under different equilibrium selection rules and examines their effect on the parameter estimates.

[7]Any mutually beneficial deviation is self-enforcing unless it is vulnerable to a self-enforcing deviation by a proper subset of the players who deviated in the first place (and so on). This internal consistency requirement distinguishes CPNE from Robert Aumann's (1959) *strong equilibrium* which must not be vulnerable to deviations by *any* sub-coalition, including those that are not robust to further deviations.

[8]Bernheim et al. (1987) define *perfectly coalition-proof Nash equilibrium* for games in extensive form, but its recursive definition limits the applicability of this concept to the study of finite-horizon games (Bernheim and Ray, 1989).

[9]The definitions by Farrell and Maskin (1989a), Bernheim and Ray (1989) and Pearce (1987) are among the most frequently cited. See also Pearce (1990) for a survey and discussion of the literature on renegotiation-proofness in repeated games.

Definition 4 (Strongly Renegotiation-Proof Equilibrium) *(Farrell and Maskin, 1989a) A WRP equilibrium is* strongly renegotiation-proof (SRP) *if none of its continuation equilibria is strictly Pareto-dominated by another WRP equilibrium.*

In essence, weak renegotiation-proofness imposes collective rationality along the equilibrium path: if any two continuation equilibria are Pareto-ranked, the dominated equilibrium is ruled out and players collectively choose the better one. Strong renegotiation-proofness imposes collective rationality in a more comprehensive fashion by ruling out WRP equilibria that do not satisfy an external dominance criterion.

Once collective rationality is imposed, finding the unique equilibrium of the dynamic model is straightforward: From period 2 onwards, the Pareto-efficient Nash equilibrium is full cooperation. Therefore, indefinite repetition of this equilibrium is SPNE. It is also WRP, because no other profile is played along the equilibrium path. Finally, since every action profile other than full cooperation gives a strictly lower payoff in the stage game, it cannot be played as part of a SRP continuation equilibrium. Hence, the only such equilibrium in period 2 is to repeat the fully cooperative outcome indefinitely.

The unique continuation equilibrium in period 2 serves as an end point from which to use backward induction. Since no punishment of prior deviations is possible beyond period 1, it follows that a Nash equilibrium of the stage game must be played in all earlier periods. In the case of constant differences, this gives rise to a unique sequence of stage game Nash equilibria. With increasing differences, the SRP equilibrium rules out play of the dominated Nash equilibrium in period 1. Therefore, the game has a unique SRP equilibrium. This equilibrium also satisfies the demand that strategies be robust to deviations by sub-coalitions of players. As I shall show below, SRP equilibrium in fact requires that the unique CPNE be played at each stage of the game.

The following section derives the formal solution to the game with asymmetric players. Readers who are more interested in the empirical application may skip directly to section 3.5.

3.2.2 The general case

In this section, I formally derive the non-cooperative solution under the more general assumptions made above. To begin, I establish existence of a pure-strategy Nash equilibrium in the stage game and identify several sources of multiplicity of (subgame-perfect) Nash equilibrium. Next, I characterize the unique SRP equilibrium of the game G with and without strategic complementarities. All proofs are in appendix A.

Existence of a Nash equilibrium in pure strategies is a well-established result in the theory of supermodular games (Topkis, 1998).

Theorem 1 *The stage game $\Gamma(t)$ has a Nash equilibrium in pure strategies for all $t \in T$.*

Somewhat trivially, uniqueness of Nash equilibrium may fail, if a player i is indifferent between two actions in some period $t' \in T$. In this case, player i's best response

$$R(a_{-i}, t) \equiv \arg\max_{a_i \in A_i} \pi_i(a_i, a_{-i}, t). \tag{3.3}$$

is the set $\{0, 1\}$, which implies that there are at least two Nash equilibria in period t'. I shall rule out this possibility by assuming that players are "benign", in the sense that they choose to cooperate whenever they are indifferent between both actions.[10]

Assumption 7 (Benign players) $\forall t \in T, \forall i \in I$, $\forall a^*_{-i} \in A_{-i}$ *such that* $\pi_i(1, a^*_{-i}, t) = \pi_i(0, a^*_{-i}, t) : \; R_i(a^*_{-i}, t) = 1$.

Equilibrium with constant differences Consider first the game in which the relative payoff to cooperation is invariant with respect to the actions of other players.

Theorem 2 *Under assumptions 1, 3 and 7, for all $t \in T$, the game $\Gamma(t)$ has a unique Nash equilibrium in dominant strategies given by*

$$a^*(t) = \big(a^*_i, i \in I | a^*_i = \mathbb{1}\{\pi_i(1, 0, t) \geq \pi_i(0, 0, t)\}\big).$$

The Nash equilibrium $a^*(t)$ is weakly increasing in t.

While a formal proof of the theorem is given in appendix A, the intuition behind it is immediate from figure 1a, where the relative payoff to cooperation is invariant to other players' actions and shifts monotonically over time.

Consider now the subgame-perfect Nash equilibrium of the infinite-horizon game which is formed by the sequence of stage-game Nash equilibria. The monotonicity of Nash equilibrium with respect to t implies that a player whose dominant action is to cooperate in a given period will never want to change her action in any period after that. Therefore, player i's provision

[10] This assumption is not very restrictive for two reasons. First, the equilibria it rules out in the dynamic game differ from the chosen equilibrium in just a finite number of periods. This is because the strict monotonicity assumption 1 guarantees that potential ties in payoffs cannot last for longer than one period. Second, in the empirical framework such ties occur with probability zero since I introduce continuous random disturbances in the payoff function.

time is determined by the first period in which cooperation becomes a dominant action:

$$t_i^0 \equiv \{t \in T | \Delta\pi_i(0, t-1) < 0 \leq \Delta\pi_i(0, t)\}. \tag{3.4}$$

Without loss of generality, it is assumed that subscripts are consistent with the ordering of t_i^0 so that

$$t_1^0 \leq t_2^0 \leq \cdots \leq t_N^0.$$

It bears noting that this SPNE is not unique. By the folk theorem, any feasible and individually rational payoff vector can be implemented as a SPNE of the continuation game $G(t_N^0)$ if players are sufficiently patient. Hence, the concept fails to rule out equilibria in which players play dominated actions infinitely often. This outcome seems rather implausible, because full cooperation is the unique Nash equilibrium and strictly Pareto-dominates all other action profiles in every stage game from t_N^0 on. Dominated SPNE are ruled out by the requirement that the outcome not be vulnerable to renegotiation by all players at any stage of the game.

Theorem 3 *Under assumptions 1-3 and 5-7 the strategy profile $s^* = \{s^*(t)\}_{t \in T}$ where*

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

is the unique strongly renegotiation-proof equilibrium of the game G, in which each player i starts cooperating in period t_i^0 and never defects at any node after that.

The theorem states that, once the SRP criterion is imposed, a single SPNE remains in which the unique stage-game NE is played in every period.[11]

Equilibrium with increasing differences Consider now the case where the relative payoff to cooperation is increasing in the number of other players who cooperate, as implied by assumption 4. This particular form of payoff interdependence renders cooperation a *strategic complement* (Bulow et al., 1985), in the sense that the incentive to cooperate is higher the more other players cooperate.

Existence of a stage-game Nash equilibrium in pure strategies in every period follows from theorem 1. As shown in figure 1b, the equilibrium may not be unique in the presence

[11] As demonstrated by van Damme (1989) and Farrell and Maskin (1989b) for the infinitely repeated prisoner's dilemma, the restrictions on the set of admissible equilibrium payoffs and discount factors imposed by weak renegotiation-proofness on the one hand and SPNE on the other hand are identical in certain games. In the present game, WRP equilibrium sustains a large set of payoff vectors and hence is not helpful in reducing multiplicity of equilibrium.

of strategic complementarities. Another way of seeing this is by inspection of assumption 4 for the profiles $a_{-i} = \mathbf{0}$ and $a'_{-i} = \mathbf{1}$: $\Delta\pi_i\left(\mathbf{1}, t\right) \geq \Delta\pi_i\left(\mathbf{0}, t\right)$ $\forall i \in I$. No assumption has been made that would rule out the existence of some $t' \in T$ such that $\Delta\pi_i\left(\mathbf{1}, t'\right) \geq 0 > \Delta\pi_i\left(\mathbf{0}, t'\right)$ $\forall i \in I$. Hence, full cooperation and no cooperation constitute the greatest and the least Nash equilibrium of the stage game $\Gamma(t')$, respectively. If players are asymmetric, there may be additional Nash equilibria.

An implication for the solution of the dynamic game G is that more than one SPNE can be constructed as sequences of stage-game NE constitutes. The refinement of strong renegotiation-proofness rules out any such equilibrium that induces a dominated Nash equilibrium in any period. The SRP equilibrium of the game can thus be found by computing the largest Nash equilibrium for every stage of the game. To this end, I devise an algorithm that starts in the first period in which full cooperation is a Nash equilibrium and, working backwards in time, drops players from the set of contributors only if they have a incentive to defect when the actions of all other players are held fixed. Since the Nash equilibrium monotonically increases over time, this procedure is guaranteed to find the largest Nash equilibrium in every period.

In order to prove this, it is convenient to write the payoff alternately as a function of an action profile $a \in A$ or – in slight abuse of notation – as a function of the set $K(a) \subseteq I$ of cooperating players induced by that profile,

$$\pi_i(K(a), t) \equiv \pi_i(a, t)$$

where $K(a) = \{i \in I : a_i = 1\}$, $K(a_{-i}) = \{j \in I, j \neq i : a_j = 1\}$ and

$$\Delta\pi_i(K(a_{-i}), t) \equiv \pi_i(1, a_{-i}, t) - \pi_i(0, a_{-i}, t).$$

Define $t_{(m)} \in T$ as the earliest period in which cooperation by a set of m players, K_m, is a Nash equilibrium of the stage game. This implies that

$$\Delta\pi_i\left(K_m \backslash \{i\}, t_{(m)}\right) \geq 0 \qquad \forall i \in K_m \tag{3.5}$$

$$\wedge \qquad \Delta\pi_j\left(K_m \backslash \{j\}, t_{(m)}\right) < 0 \qquad \forall j \in I \backslash K_m \tag{3.6}$$

Indexing by subscript (m) the player $j_{(m)} \in K_m$ with the least incentive to cooperate, we have that

$$\Delta\pi_{(m)}\left(K_m \backslash \{j_{(m)}\}, t_{(m)}\right) \leq \Delta\pi_i\left(K_m \backslash \{i\}, t_{(m)}\right) \qquad \forall i \in K_m \tag{3.7}$$

$$\Delta\pi_{(m)}\left(K_m \backslash \{j_{(m)}\}, t_{(m)} - 1\right) < 0. \tag{3.8}$$

For $m = N, N-1, \ldots, 1$, eqs. (3.5)-(3.8) recursively define N sets $K_{(m-1)} \equiv K_m \backslash \{j_{(m)}\}$, the N action profiles inducing them, $a^m \equiv (a_j, j \in I | a_j = \mathbb{1}\{j \in K_m\})$, and N periods $t_{(m)}$. Using this notation, the SRP equilibrium can be characterized as follows.

Theorem 4 *Under assumptions 1-2 and 4-7 the game G has a unique strongly renegotiation-proof Nash equilibrium $s^* = \{s^*(t)\}_{t=0}^{\infty}$ where*

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

Player $j_{(i)}$ starts to cooperate in period $t^*_{(i)}$ and never defects at any time after that. The sequence $\{t^*_{(i)}\}_{i=1}^{N}$ is given by $t^*_{(N)} = t_{(N)}$ and

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

where $t_{(m)}$ is defined in eq. (3.5)-(3.8) *for all $m = 1, \ldots, N$.*

The effects of strategic complementarities on equilibrium provision become apparent when comparing the outcomes characterized in theorems 3 and 4: First, they bring forward public good provision in time by reinforcing the incentive to cooperate. Second, if this effect is sufficiently strong, cooperation by one player triggers cooperation by one or more other players. In section 3.5 below, I shall discuss how this model prediction can be exploited in an empirical application to learn about the strength of strategic complementarities.

In the discussion of appropriate equilibrium concepts in section 3.1.3 above, it was mentioned that it would be desirable to find an outcome to deviations by proper subsets of players of arbitrary size. Farrell and Maskin (1989a) note that "renegotiation-proofness is to some extent a *cooperative* requirement" (p. 355) in games with more than two players, because it disregards deviations by fewer than all players. In the present game, such deviations are not much of a concern because all Nash equilibria are Pareto-ranked. As a consequence, the Nash equilibrium induced by SRP equilibrium at every node of the game coincides with the unique coalition-proof Nash equilibrium.

Theorem 5 *The largest Nash equilibrium in every stage game $\Gamma(t)$, $t \in T$ corresponds to the unique coalition-proof Nash equilibrium of $\Gamma(t)$.*

3.3 Socially Optimal Provision

It has been assumed thus far that players behave in a non-cooperative fashion. Plainly, the concept of Nash equilibrium rules out levels of public good provision that are not individually rational, and the refinement of SRP imposes collective rationality only to the extent that it does not run counter to this basic premise. The non-cooperative game theoretical framework is deemed the most appropriate for the analysis of public good provision in the absence of binding agreements and external enforcement.

Nonetheless, it is instructive to compare the outcomes obtained under that assumption with the case of a benevolent planner who chooses players' actions in each period so as to maximize the cumulative discounted sum of payoffs across all players. Write aggregate payoffs in period t given profile $a(t)$ as

$$W(a(t),t) = \sum_{i=1}^{N} \pi_i(a(t),t) \tag{3.9}$$

and the planner's decision problem as

$$\max_{a(t)\in A} \sum_{t=0}^{\infty} W(a(t),t)e^{-\rho t} \tag{3.10}$$

where ρ is the social discount rate. Since choices in period t do not have dynamic repercussions this problem is equivalent to

$$\max_{a\in A} \sum_{j=1}^{N} \pi_j(a,t) \qquad \forall t \in T. \tag{3.11}$$

The solution to problem (3.11) is given by a sequence of optimal action profiles $\{a^o(t)\}_{t\in T}$.

Theorem 6 *For any period $t \in T$, if a^o is a maximizer of $W(a,t)$ in A and a^* is a Nash equilibrium of the stage game $\Gamma(t)$, then $a^o \geq a^*$.*

Theorem 6 implies that a social planner will have all players contributing to the public good no later than they do in the non-cooperative scenario. This means that aggregate welfare would increase if some players could be coerced to cooperate earlier. This is reminiscent of the prisoners' dilemma where full cooperation maximizes aggregate payoffs. As a useful byproduct, theorem 6 helps with computation of the solution to the planner's problem by reducing the number of iterations in a grid search for the global optimum of $W(a,t)$. In particular, the result tells us that the search can be constrained to the set of strategy profiles larger or equal to the largest Nash equilibrium $a^*(t)$ of the stage game, i.e.

$$\arg\max W(a,t) \subseteq \{a \in A : a \geq a^*(t) \vee a = a^*\}.$$

3.4 Limit of the Discrete-Time Game

Here I examine the limit of the equilibrium outcomes of the game when time periods become infinitesimally small. Consider first the continuous-time (or calendar time) analogues to the equilibrium provision times in the discrete-time game. The first calendar time at which cooperation is a dominant strategy for player i is given by $\tilde{t}_i^0 = \{\tilde{t} \in \mathbb{R}_+ | \Delta\pi_i(0,\tilde{t}) = 0\}$. Moreover, the first calendar time at which full cooperation is a Nash equilibrium of the stage game is given by $\tilde{t}^*_{(N)} = \min\{\tilde{t} \in \mathbb{R} | \Delta\pi_i(I\backslash\{i\},\tilde{t}) \geq 0 \, \forall i \in I\}$. The calendar times corresponding to SRP equilibrium in the game with strategic complementarities can be computed in an analogous fashion to the algorithm described in section 3.2.2.[12]

At the heart of the definition of provision times both with and without strategic complementarities is the condition

$$\Delta\pi_i(K,t) \geq 0 \tag{3.15}$$

where K is some set of contributors; for example, $K = \emptyset$ pins down t_i^0 and $K = I\backslash\{i\}$ pins down $t^*_{(N)}$. The only difference between $t^* \in T$ and $\tilde{t}^* \in \mathbb{R}_+$ is that the latter solves this condition with equality (due to the continuity assumption 1 and by the intermediate value theorem) whereas in the discrete-time game one needs to find the first period on the grid satisfying the weak inequality. The relationship between the two on a grid with period length one 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}, $k \in \mathbb{N}_0$. A node on this grid, ${}^kt$, has the property ${}^kt = t2^k$. That

[12] Denote by subscript m the player $j_{(m)} \in I$ who is just indifferent between cooperating or not in the stage game played at time $\tilde{t}_{(m)}$ if all players in K_m cooperate. Compute this time as

$$\tilde{t}_{(m)} = \{\tilde{t} \in \mathbb{R} | \Delta\pi_{(m)}(K_m\backslash\{j_{(m)}\},\tilde{t}_{(m)}) = 0 \\ \wedge \; \Delta\pi_i(K_m\backslash\{i\},\tilde{t}_{(m)}) \geq 0 \; i \in K_m \; \wedge \Delta\pi_j(K_m\backslash\{j\},\tilde{t}_{(m)}) < 0 \; j \notin K_m\} \tag{3.12}$$

Define

$$K_{m-1} = K_m\backslash\{m\}. \tag{3.13}$$

Let $K_m = I$ and compute $\tilde{t}_{(m)}, j_{(m)}, K_{(m)}$ by recursive application of (3.12) and (3.13) for $m = N, N-1, \ldots, 1$. The equilibrium provision times are given by $\tilde{t}^*_{(N)} = \tilde{t}_{(N)}$ and

$$\tilde{t}^*_{(N-i)} = \min\left[\tilde{t}_{(N-i+1)}, \tilde{t}_{(N-i)}\right] \qquad i = 1, \ldots N-1. \tag{3.14}$$

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. Equilibrium in this limit is characterized in the following theorem.

Theorem 7 *As time periods become infinitesimally small, the equilibrium of the discrete-time game ${}^{\infty}G$ exists and equilibrium provision times converge to their continuous-time analogues,*

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

3.5 Implications for Empirical Analysis

This section derives restrictions that can be exploited for the empirical estimation of the timing model of public good provision using data on binary provision decisions over time. It is assumed that the researcher observes at least one history of the game with a large number of players N. The researcher is interested in estimates of both the private net benefits of cooperation and the magnitude of strategic complementarities. The following assumption puts some structure on the way heterogeneity enters the relative payoff to cooperation.

Assumption 8 (Payoff heterogeneity) $\forall i \in I, \tilde{t} \in \mathbb{R}$, $a_{-i} \in A_{-i}$;

$$\Delta\pi_i(a_{-i}, \tilde{t}) \equiv r(\phi_i, a_{-i}, \tilde{t})$$

where $\phi_i \in \mathbb{R}$ is a continuous, i.i.d. random variable with distribution function $F(\phi)$ and r is a continuously differentiable and strictly monotonic function of ϕ_i:

$$r(\phi_i, a_{-i}, \tilde{t}) \in \mathrm{C}^1 \;\wedge\; \frac{\partial r(\phi_i, a_{-i}, \tilde{t})}{\partial \phi_i} \neq 0.$$

Given these assumptions, if the payoff function has constant differences then clustering of provision times is a probability-zero event. This is the key insight underlying the following theorem.

Theorem 8 *If assumptions 1, 2 and 5-8 hold and time periods become infinitesimally small, then clustering of provision decisions among asymmetric players occurs only in the presence of strategic complementarities.*

The theorem suggests a strategy to empirically identify strategic complementarities in the timing of public good provision. The basic intuition is as follows: Given a data set in which individuals start out defecting and cooperate only at a later stage, the timing of the decision to cooperate traces out the distribution of the relative benefits to cooperation.

Furthermore, if provision times are recorded on a sufficiently fine time grid, clustering must be attributed to strategic complementarities unless players are identical.

This identification strategy rests on two key assumptions that I briefly discuss here. First, assumption 8 guarantees that players are not identical with probability one by introducing a random disturbance that shifts the payoff function in a strictly monotonic fashion. This assumption is very common in applied research and prevents "overfitting" of the model.

A more restrictive assumption is the assumption of a smooth time trend in $\Delta\pi$ (assumption 1). If the relative payoff to cooperation were allowed to jump then two or more players who are close to the threshold before the jump will start to cooperate immediately after the jump even as the grid length goes to zero. This leads to an overestimation of the magnitude of strategic complementarities. Conversely, a sudden drop in relative benefits to cooperation would result in underestimation of strategic complementarities. Assumption 1 rules out such jumps. It bears noting, that the assumption is not more restrictive than the assumption of a hazard rate that is constant or a smooth function of time, which is commonly made in duration analysis. In contrast to a duration model, however, the present framework explicitly accounts for forward-looking strategic behavior of players who anticipate mutually beneficial deviations from dominated stage-game Nash equilibrium.

4 Stratospheric Ozone Depletion and the Montreal Protocol

This section provides a review of the scientific underpinnings of stratospheric ozone depletion to the extent that they bear relevance to the economic modelling of the problem. Furthermore, it summarizes the diplomatic efforts that lead to the signing of the Montreal Protocol in September 1987 and subsequent amendments. The section draws on the comprehensive accounts provided by Benedick (1998) and Parson (2003).

4.1 Background on Ozone Depletion

Chlorofluorocarbons (CFCs)[13] were invented by chemists at DuPont and General Motors in 1928, and first commercially used as working fluids in refrigerators. Due to their stable chemical structure and low production cost, CFCs were soon used as freezing agents in air conditioning, as aerosol propellants in spray cans, as plastic-foam-blowing agents and as

[13]The most common CFCs are CCl3F (CFC-11), CCl_2F_2 (CFC-12), CCl_2FCClF_2 (CFC-13), $CClF_2CClF_2$ (CFC-14), and $CClF_2CCl_2F$ (CFC-113).

solvents in the manufacturing of microchips and telecommunications parts. Along with the expansion of applications went a rapid growth in worldwide production. According to data from the Alternative Fluorocarbons Environmental Acceptability Study (AFEAS, 2006), shown in table 1, worldwide production of CFC-11 and CFC-12 between 1931 and 1977 grew on average 20.6% and 14.3% p.a., respectively. This corresponds to a doubling of CFC-11 (CFC-12) output in less than every four (five) years.

In the early 1970's, scientists began to articulate concern about the detrimental effects of CFC releases on stratospheric ozone. A new theory by chemists Paul Crutzen, Mario Molina and Sherwood Rowland[14] fundamentally challenged the prevailing view at the time that CFCs were a safe and environmentally friendly alternative to the easily inflammable and often toxic substances they were replacing in many applications. According to the theory, CFC molecules remain intact for several decades after being released into the atmosphere, helped by their exceptional stability. They are slowly being transported up to the stratosphere where solar radiation eventually breaks them apart. The chlorine atoms released in this process enter in a catalytic chain reaction with oxygen (O) atoms in ozone molecules (O_3) which interferes with the natural, solar-powered conversion cycle between ozone and oxygen. Molina and Rowland (1974) predicted that, as a consequence of unaltered CFC emissions, ozone concentrations in the stratosphere would fall substantially. This conclusion was alarming, since it was already known that stratospheric ozone acts as a shield against UV B radiation from space that causes skin cancer, eye cataracts and immune disorders as well as damage to crops and ecosystems.

The chemistry behind the theory of ozone depletion was consistent, yet scientists struggled for more than a decade to gather conclusive statistical evidence of a loss in ozone levels at all.[15] The scientific debate reached a turning point in 1985, when the British Antarctic Survey (BAS) published measurement data showing a seasonal decline in ozone concentrations between September and November of about 50% compared to the 1960s. This induced NASA scientists to revisit measurement data collected by satellite instruments which had

[14]The three researchers later shared the 1995 Nobel Prize in Chemistry for their theory of stratospheric ozone depletion. The theory was put forth by Molina and Rowland (1974) and drew on research by Crutzen (1970), who had established that nitrogen could deplete stratospheric ozone in a catalytic reaction. It also built on Stolarski and Cicerone's (1974) finding that chlorine could have the same effect, and on evidence that CFCs mix uniformly in the atmosphere (first published by Lovelock et al., 1973)

[15]Atmospheric measurements early on confirmed the presence of chlorine, hydrogen chloride (HCl) and the catalytic intermediary, ClO. However, there was substantial uncertainty about the actual magnitude of ozone depletion. Seasonal and cyclical fluctuations in stratospheric ozone levels made it difficult to measure a trend. In ever more complex modeling studies published between 1974 and 1983, estimates of global average depletion 50 to 100 years into the future ranged from 3% to 20%. As Benedick (1998, p. 13) notes, "these swings began to affect the credibility of the science and to dampen both public and official concern about the urgency of the problem."

been flagged as erroneous since they lay outside the programmed error bounds. Not only did NASA confirm the ozone loss over the Antarctic but it also reported large ozone losses worldwide.

The discovery of what would become known as the "Ozone Hole" above the Antarctic came as a shock to the public. Yet it did not by itself rule out alternative explanations.[16] It was not until 1988 that the Ozone Trends Panel, an international expert panel organized by NASA, had gathered enough evidence to refute these hypotheses. In its report, the panel concluded that worldwide ozone losses were wholly or in part due to CFCs and roughly twice as high as predicted by current models.

4.2 Ozone Diplomacy

As the scientific debate about the role of CFC in stratospheric ozone depletion gained momentum, a few countries began to implement domestic policy measures aimed at curbing CFC emissions.[17] However, policy-makers quickly came to realize that an effective solution to the ozone problem necessitated international policy coordination. After all, the science of ozone depletion indicated that it was a pure global public bad, and the fact that CFCs were being produced in many different countries jeopardized the effectiveness of unilateral regulations.

The Vienna Convention An international diplomatic effort was launched by the United Nations Environmental Programme (UNEP) with the funding of a conference on the implications of ozone research, organized by the World Meteorological Organization (WMO) in 1975. Over the following years UNEP continued to sponsor scientific research and disseminate the results. The question of international controls over CFCs was formally raised for the first time at an intergovernmental meeting in 1977, but no consensus was reached. UNEP approved a non-binding resolution in 1980 that called upon countries to reduce their use of CFC but did not specify quantitative targets. A year later, the agency began work

[16] For instance, it was hypothesized that the Antarctic ozone hole was a meteorological phenomenon unrelated to CFC release or that the measured decline in worldwide ozone concentrations could be entirely driven by instrument degradation.

[17] In 1978, the US implemented a ban on non-essential aerosol uses of CFCs – essentially in spray cans – that was followed by a sharp decline in aerosol production by 95%. The ban came after US producers had found a more economical substitute for CFC use in aerosol propellant.While Canada, Norway and Sweden adopted similar bans, the European Parliament rejected calls for a ban in the European Community (EC). Only in 1980 did the EC member countries agree to curb CFC aerosol use by 30% of 1976 levels until the end of 1981. Moreover, the EC capped production capacity for CFC-11 and CFC-12 at its 1980 level and later promoted measures to conserve CFCs and to inhibit their release into the atmosphere. Benedick (1998, p.25, p.41) judges the caps on production and capacity as "trivial" and "inconsequential" because CFC emissions had already fallen by 28% at the time and utilization rates were low.

on an international agreement for protecting the ozone layer and continued persistently over the next few years, even as political interest in the issue declined. These efforts culminated in March 1985 with the signing of the Vienna Convention for the Protection of the Ozone Layer. Apart from extensive scientific research on ozone depletion, the treaty formulated a general obligation for states to "adopt appropriate legislative or administrative measures" to protect the ozone layer – though it did not specify particular measures such as targets and timetables for emission reductions. Importantly, the convention set up the Conference of Parties, an institution working toward a future protocol to control emissions of ozone depleting substances.

The Montreal Protocol The first protocol under the Vienna Convention was opened for signature two years later, in September 1987. The treaty, called the Montreal Protocol on Substances that Deplete the Ozone Layer, stipulated that a 50% cutback in consumption of five CFCs (CFC-11, CFC-12, CFC-113, CFC-114, CFC-115) from 1986 levels be achieved in three stages: a stabilization in 1989 and reductions by 20% and 50% by 1993 and 1998, respectively. 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. Addressing concerns about leakage, the treaty prohibited bulk imports of controlled substances from non-member states (in 1990) and bulk exports from developing countries (in 1993). Trade between developed member countries was left unrestricted, but imports were counted towards a country's net consumption.[18] Imports of products containing controlled substances were banned from 1992 and the treaty stipulated that parties inquire into the feasibility of banning or restricting imports of products not containing, but manufactured with controlled substances. Member states were also required to provide data on production, consumption and trade of controlled substances, and to promote technical assistance to facilitate participation of developing countries and implementation. Finally, parties agreed to hold regular meetings to review the implementation of the treaty and make necessary adjustments to its regulations.

The Montreal Protocol went into force on January 1, 1989, by which time it had been ratified by most major developed countries. Remarkably, the first country to ratify the treaty was Mexico, a developing country. By contrast, other large developing countries, such as India and China, showed no intentions of ratifying. Out of concern about leakage of CFC production to non-member states, amendments were made to the original treaty during the

[18]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".

second meeting of parties held in London in June 1990. The amendments established a multilateral fund to pay for the incremental costs developing country parties faced in complying with their obligations. Contributions to the fund were the responsibility of developed country parties according to the UN scale of assessment. This financial incentive for accession was accompanied by the promise of making technology available at favorable terms.

The time table for phasing out ozone depleting substances was brought forward at several meetings. During the London negotiations, industrialized country parties agreed to a complete phase-out of CFCs by 2000, with interim goals of 50% in 1995 and 85% in 1997. Two years later, the meeting of parties in Copenhagen revised the goals for CFC to a 75% reduction in 1994 and a phase-out by 1996. The meeting set the deadline for the phase-out of halons by 1994 and provided a time table for the phase-out of HCFCs – an early substitute for CFCs later found to be ozone depleting – by 2030. The schedule for developing countries remained unchanged, stipulating the phase-out of most controlled substances by 2010. Further meetings of parties at Vienna (1995), Montreal (1997) and Beijing (1999) augmented the list of controlled substances to currently include 98 items and adopted measures to address issues such as illegal trade. The global phase-out of CFCs is proceeding swiftly within the parameters set by the Montreal Protocol, making the treaty one of the most successful international environmental agreements in history.

5 An Empirical Model of Treaty Ratification

5.1 Data

A key issue in empirical work on international environmental agreements is the lack of data on a cross-section of independent agreements. Consequently, it is not feasible to estimate a model that merely predicts the equilibrium number of participants, such as the model of self-enforcing international agreements, because an agreement provides just a single observation for estimation. Pooling observations across different treaties does not solve this problem, because countries' net benefits will typically differ widely across treaties.

This paper proposes a different empirical approach that exploits the variation in the timing of ratification decisions. Figure 2 plots the time to ratification of the Montreal Procotol on Substances that Deplete the Ozone Layer – measured in days after it was opened for signature on September 16, 1987 – against the natural log of population (left) and log per capita GDP in 1986 (right, expressed in 1995 US$). The figure suggests that there is a negative association between per capita income and the ratification time. Moreover, while ratification by some countries is spaced out over time, others ratify the treaty in

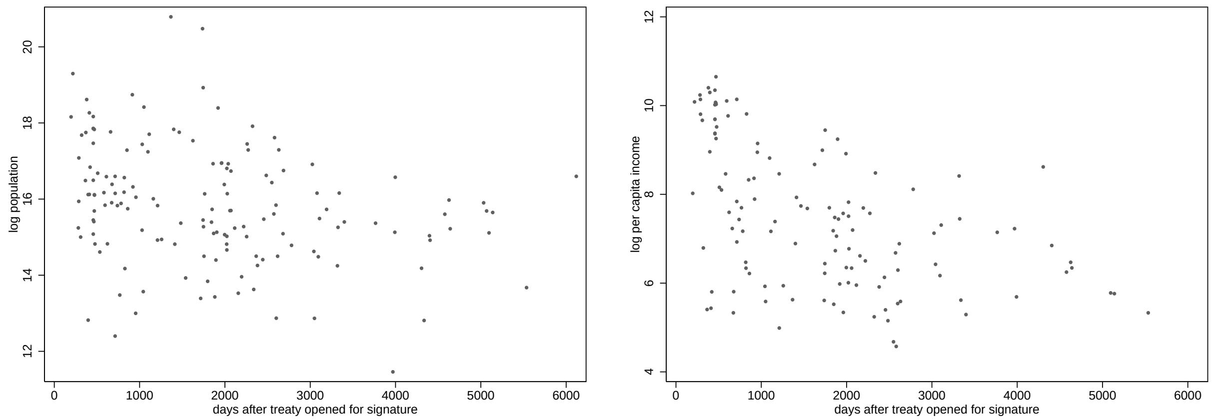


Figure 2: The Timing of Ratification

bunches, e.g. in the same month or week, or even on the same day. Anecdotal evidence suggests that this bunching is not by coincidence. For example, the EC commission sought to coordinate ratification by all EC member states on a single date (Benedick, 1998, p. 161). The estimation strategy adopted below is based on the notion that the variation in the timing of ratification contains relevant information about strategic interaction between countries. It adapts the general model presented above to the issue of ozone depletion and exploits its predictions about clustering and strategic complementarities. The goal is to estimate the magnitude of endogenous reinforcement effects in the ratification of the Montreal Protocol.

5.2 Model Primitives

Participation in the Montreal Protocol is modeled as an infinite-horizon game played among N countries. In each period $t = 0, 1, \dots$ all countries choose simultaneously whether to join (remain in) the treaty or to stay outside (withdraw). Players are assumed to have complete information about all primitives of the game.

Consider the following per-period payoff function for country i:

$$\pi_i(a_i(t), a_{-i}(t), t) = \begin{cases} EB_i(a_{-i}(t), t) & \text{if } a_i = 0 \\ NB_i(t) + EB_i(a_{-i}(t), t) + S_i(a_{-i}(t)) & \text{if } a_i = 1. \end{cases} \tag{5.1}$$

Country i's private net benefit of joining the treaty is assumed to take the form

$$NB_i(t) = -\phi_i + e^{\lambda t}. \tag{5.2}$$

This specification accounts for the well-known fact that discrete choice models do not allow for separate identification of the private costs and benefits associated with an action. The term ϕ_i denotes an exponential index of net cost

$$\phi_i = \exp(x_i'\beta + \epsilon_i) \tag{5.3}$$

where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon)$ is a vector of independently and identically distributed error terms that are unobserved by the econometrician but observed by all players. The vector x_i contains country characteristics that shift the net benefit of ratification.

The net benefit is assumed to be exponentially increasing at the rate $\lambda > 0$. This reflects the fact that the marginal damage of CFC emissions arises from an exponentially growing stock of CFC pollution in the stratosphere. As shown in the first column of table 1, worldwide CFC emissions grew at double-digit rates between 1931 and 1977. Since CFCs travel slowly and chlorine is released only gradually, ozone depletion during the late 1980's and 1990's was caused by CFC emitted between the 1930's and the 1950's (see the discussion in section 4). The evolution of the marginal benefit to abatement is taken as exogenous because contemporaneous abatement under the Montreal Protocol does not affect current ozone depletion.

Scope for strategic interaction arises from two types of spillovers. First, since ozone depletion is a global public bad, each country enjoys the external benefits of other countries' abatement, $EB_i(a_{-i}(t), t)$, regardless of its own action. $EB_i(a_{-i}(t), t)$ is assumed to be strictly increasing in a_{-i}. Second, I allow for a spillover

$$S_i(a_{-i}(t), t) = \gamma \textstyle\sum_{j \neq i} w_{ij} a_j(t) \tag{5.4}$$

that is exclusive to treaty members. The term $w_{ij} \geq 0$ denotes the intensity of the spillover from j to i, and $\gamma \geq 0$ is a parameter. A positive γ makes ratification of the treaty a strategic complement. Strategic complementarities between treaty members may arise because of the trade ban in controlled substances, due to economic dependencies, out of concern about reputation and equity, and from other kinds of social interaction. In section 7 I shall investigate the sources of such effects further by calibrating the terms w_{ij} on observable spillovers.

5.3 Equilibrium

Country i's optimal strategy s_i is the solution to

$$\max_{s_i \in S_i} V_i(s_i, s_{-i}, t) = \sum_{\tau=t}^{\infty} \pi_i(s_i(\tau), s_{-i}(\tau), \tau) e^{-r_i(\tau - t)} \tag{5.5}$$

where $\pi(\cdot)$ is given by eq. (5.1) and given a strategy profile of all countries other than i, s_{-i}. The strongly renegotiation-proof Nash equilibrium of the model is found by straightforward application of the results obtained in section 3. Equilibrium ratification times are determined by the relative payoff to cooperation

$$\Delta\pi_i(a_{-i}, t) = -\phi_i + e^{\lambda t} + \gamma \textstyle\sum_{j \neq i} w_{ij} a_j(t) \tag{5.6}$$

and depend on the private net benefit ϕ_i, on the rate at which the benefit increases λ, and on the strength of the strategic complementarity γ.

Theorem 9 *The game has a unique strongly renegotiation-proof Nash equilibrium in which each country i joins the agreement at time $t^*_{(i)}$ and never withdraws after that. The sequence of ratification times $\{t^*_{(i)}\}_{i \in I}$ can be computed as follows:*

1. *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*

$$\tilde{t}_i(K) = \frac{1}{\lambda} \log \left[\phi_i - \gamma \left(\sum_{j \in K \setminus \{i\}} w_{ij} \right) \right] \quad , i \in K. \tag{5.7}$$

   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 $t^*_{(N)} = \lceil \tilde{t}_{(N)}(I) \rceil$.

2. *Recursively, for $m = N-1, N-2, \ldots, 1$, define the set $K_m \equiv K_{m+1} \setminus j_{(m+1)}$ 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 theorem implies that, if $\gamma = 0$, a country's optimal ratification time is given by $t^0_i = \lceil \frac{1}{\lambda} \ln(\phi_i) \rceil$, sweeping out the distribution of ϕ_i. In the presence of strategic complementarities, the timing of ratification is also determined by the magnitude of the spillovers γw_{ij} between any two countries i and j.

5.4 Parameter Identification

The vector of model parameters to be estimated is given by (γ, β, λ). Notice that the external benefits EB_i are not identified since they enter the payoffs for both actions in the same fashion. Likewise, the discount rate r_i has no bearing on the solution of the model. As is common in discrete choice models, the variance σ_ϵ is not identified. For fixed σ_ϵ, 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.

There are two different sources of identification of the spillover parameter γ. One is the clustering of ratification times and the other is the spillover intensity w_{ij} which may or may not be observed. Below, I introduce more specific sources of identification for the coefficient γ by considering three different channels through which spillovers might work. Regarding the former source, theorem 8 implies that γ is parametrically identified provided that some clustering is observed in the data. This identification strategy critically depends on the assumption that it is not common shocks to ϕ that generate the clustering observed in the data. For example, unanticipated drops in the cost of CFC abatement could push several countries over the threshold and cause clustering of ratification decisions at the time the cost drop is realized. The empirical model would interpret this as $\gamma > 0$. By contrast, a gradual and smooth decline in the cost of CFC substitutes would get absorbed into the estimate of λ and leave the estimate of γ unaffected. This is because the empirical model cannot discriminate between smooth reductions in costs and smooth growth in the benefits of abatement over time. Finally, a cost shock that occurs in a single country would get absorbed into the estimate of ϕ and leave γ unaffected because it would not cause clustering in the ratification.

A study by Hammitt (2000) finds that the time path of abatement cost did not exhibit large and discrete jumps, suggesting that a possible confounding of endogenous effects with cost reductions is not a concern. Hammitt studies the prices of CFC allowances that were traded under a cap and trade scheme imposed in the USA between 1989 and 1995. He explains that emission caps for total CFCs – weighted by their ozone-depleting potential (ODP) – were not binding over this period while gradual increases in an excise tax shifted the demand. He further argues that this shift in demand traces out the marginal cost of abatement curve and plots this curve for the three major CFCs (CFC-11, CFC-12, and CFC-113). All curves are strictly increasing between 1989 and 1994 and show no evidence of drastic cost drops. While Hammitt's findings point to a gradual diffusion of technological alternatives to ozone-depleting substances, the assumption of the smooth trend remains untestable. As was pointed out in section 3.5, alternatives for the empirical modeling of the ratification process such as duration models impose equivalent assumptions *and* assume non-strategic behavior. The present model relaxes this assumption.

6 Econometric Estimation

6.1 Data and Specification

Table 3 reports descriptive statistics of the data in two samples: the sample shown in panel A contains only countries for which CFC emission data were available whereas panel B contains a larger sample with fewer country characteristics. The endogenous variable DAY is the day on which a country ratified the Montreal Protocol, which marks the legal act of acceding to the agreement. Ratification dates for all 191 member states were obtained from UNEP's ozone web site. The mean time to ratification in the large sample is 4 years, 9 months, and 23 days.[19]

The choice of covariates in x is subject to a tradeoff. On the one hand, one would like to include all relevant variables that could possibly shift a country's net benefits. On the other hand, the small size of the data set dictates a rather parsimonious specification so as to conserve degrees of freedom and to guard against multicollinearity. As covariates in the vector x I include per capita income, population size, and CFC consumption (all in logs: LNPCY, LNPOP, LNCFC, respectively), a country's latitude (LATDEG), dummy variables for being an article 5 country (ART5) or a CFC producing country (PRODUCER).

Table 2 lists all included covariates along with the rationale behind their inclusion and the expected signs of their coefficients. The most severe consequence of the depletion of the ozone layer is the increase in the risk of skin cancer. Per capita income and population are expected to outward the demand for environmental quality in per capita and absolute terms. Since stratospheric ozone is thinning more quickly at the poles and the high latitudes, the benefit of CFC abatement should be higher there. CFC consumption and hosting a CFC producer are both expected to shift the cost of CFC abatement. The article 5 dummy picks up both structural differences between developing and developed countries and the effect of the grace period granted to article 5 countries on the net benefits of abatement. Since in the London Amendments of June 30, 1990 it was decided to offer side payments to developing countries, I include a dummy variable (LONDON_ART5) for article 5 countries that joined after that date in order to control for the change in the incentives for accession.

Data on covariates were compiled from various sources. All covariates are reported at their 1986 values – the year before the treaty was opened for signature – in order to preclude

[19] see http://ozone.unep.org

simultaneity issues. Country characteristics such as GDP in 1995 US$ and population size in millions were taken from the World Development Indicators.[20] I use the variable "Consumption of ODS: Chlorofluorocarbons" taken from UNEP (2004) as a measure of CFC consumption. "Consumption" means production plus imports minus exports of controlled substances and is reported by the member states to the treaty secretariat that monitors compliance. The variable measures cumulative consumption (in metric tons) of all five CFCs that were regulated under the Montreal Protocol, weighted by their relative ozone-depleting potential. The Protocol permits the European Community (EC) to report an EC-wide average of CFC consumption instead of individual consumption by each of its member states. A few EC countries, like West Germany, have been reporting their individual consumption on a voluntary basis. The production variable "Production of ODS: Chlorofluorocarbons" from the same source is used to construct the producer dummy, which equals 1 for all countries that report positive production of any CFC. A list of countries with article 5 status is obtained from UNEP.[21] The variable LATDEG refers to the latitude of the country's capital and is obtained from the CIA World Fact Book.[22]

6.2 Method of Simulated Moments Estimation

The choice of an estimation routine is subject to the challenges posed by the complexity of the model and the limitations of the available data. Since the expectation of equilibrium ratification times does not have a closed-form solution in the presence of strategic complementarities, I employ the method of simulated moments[23] (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 can be computed as

$$\hat{\theta} = \min_{\theta} g(\theta)' W g(\theta) \tag{6.1}$$

where

$$g = \frac{1}{N} \sum_{i=1}^{N} \left(\mu(t_i) - \frac{1}{S} \sum_{s=1}^{S} \mu(t_i(\epsilon^s; \theta)) \right) \times f(X_i) \tag{6.2}$$

[20] available online at http://devdata.worldbank.org/dataonline
[21] available online at http://ozone.unep.org/
[22] available online at https://www.cia.gov/cia/publications/factbook/index.html
[23] Maximum likelihood estimation, while desirable from an efficiency point-of-view, is not feasible because of an abundance of probability-zero events.

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

$$\left(1+\frac{1}{S}\right)\left[E_0\frac{\partial g'}{\partial\theta}W^{-1}\frac{\partial g}{\partial\theta'}\right]^{-1}E_0\frac{\partial g'}{\partial\theta}W^{-1}E_0gg'W^{-1}\frac{\partial g}{\partial\theta'}\left[E_0\frac{\partial g'}{\partial\theta}W^{-1}\frac{\partial g}{\partial\theta'}\right]^{-1}. \tag{6.3}$$

Using the optimal weighting matrix $W^*=E_0gg'$ the asymptotic covariance matrix of $\sqrt{N}(\hat{\theta}-\theta)$ can be reduced to

$$\left(1+\frac{1}{S}\right)\left[E_0\frac{\partial g'}{\partial\theta}W^{*-1}\frac{\partial g}{\partial\theta'}\right]^{-1}. \tag{6.4}$$

In addition to the structural error term ϵ, I introduce a second error term $\eta\sim\mathcal{N}(0,1)$ that represents "optimization error" – i.e. random deviations from the optimal ratification times. For example, if $\tilde{t}^*(\epsilon^s;\theta)$ is the vector of optimal ratification times for a given parameter vector θ and random draw ϵ^s, the final vector of ratification dates is computed as $t(\epsilon^s,\eta^s;\theta)=\lfloor\tilde{t}^*(\epsilon^s;\theta)+\eta^s\rfloor$. The algorithm computes the simulated moments and evaluates the criterion function (6.1) with W taken to be the identity matrix. Minimization of (6.1) in this fashion yields an initial consistent estimate of θ that is used to compute a consistent estimate of W^*. Next, the estimation is repeated using the optimally weighted criterion function to obtain the final MSM estimator $\hat{\hat{\theta}}$. Standard errors are computed on the basis of eq. (6.4) using $S=1000$ random draws and the estimate $\hat{\hat{\theta}}$ to compute the Jacobian of the moment conditions. Alternatively, standard errors are computed in a non-parametric bootstrap with 150 random samples.[24]

The moments are chosen so as to capture relevant variation in the data that helps to identify the parameters of the model. I match ratification times interacted with all other covariates and with a set of dummies that splits the sample into five bins of (nearly) equal size. In order to identify the spillover effect I match the absolute difference between i's ratification time and that of j, weighted by w_{ij}; $\sum_{j\neq i}|t_i-w_{ij}t_j|$. I also match the frequency of countries ratifying on the same day, $\mathbb{1}\{t_i=t_{i+1}\}$.

For the minimization of the MSM criterion function (6.1) I sequentially employ two algorithms. First, I use Goffe et al.'s (1994) implementation of the simulated annealing algorithm. The initial temperature value is chosen in such a way that the algorithm starts

[24] In the results below, all standard errors are bootstrapped, with the exception of table 7. Bootstrapped standard errors for this specification will be available in an updated version of this paper to be posted on my web site shortly.

out with high rejection probabilities that mimic a random grid search. This is 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.

6.3 Estimation Results

6.3.1 Parameter estimates

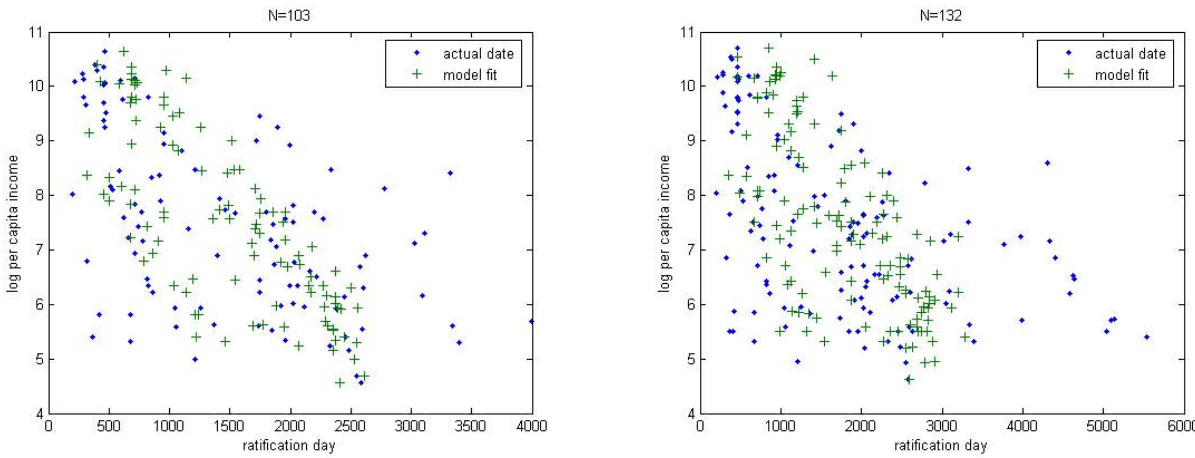
Table 4 reports estimates of the parameters in eq. (5.6) when symmetric weights $w_{ij} = (N-1)^{-1}$ are imposed. This specification serves as a benchmark against which to compare the effect of different sources of spillovers. Bootstrapped standard errors are in parentheses.

The estimate for γ is positive and significant for all specifications, indicating that ratification decisions are strategic complements. The parameters estimates for variables that shift the net cost of complying are quite robust across specifications and largely confirm the intuition summarized in table 2. 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 is consistent with the fact that more populous countries benefit more from global abatement efforts in absolute terms than smaller 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. The dummy variable LONDON_ART5 controls for changes that were made to the treaty at a later stage. Its inclusion lowers the estimated coefficient for γ and the constant term, and it flips the sign on LNCFC. The positive coefficient on CFC consumption in column 2 suggests that it is more costly for large consumers to phase out use of these substances.

In column 3, LNCFC is excluded from the empirical model which leads to slightly lower parameter estimates for LNPCY and LNPOP, both of which are positively correlated with LNCFC. The last column reports the results obtained with the larger data set that does not include CFC consumption. As evident from table 3, the additional countries on average have later ratification dates. This is reflected by a lower estimated value of λ and a slightly lower value of γ.

6.3.2 Goodness of fit

Figure 3 plots actual and fitted ratification dates from the strategic model against log per capita income for both samples. The plots (a) and (b) are based on the parameter estimates reported in columns (2) and (4) of table 4, respectively. Overall, the model fits the data quite



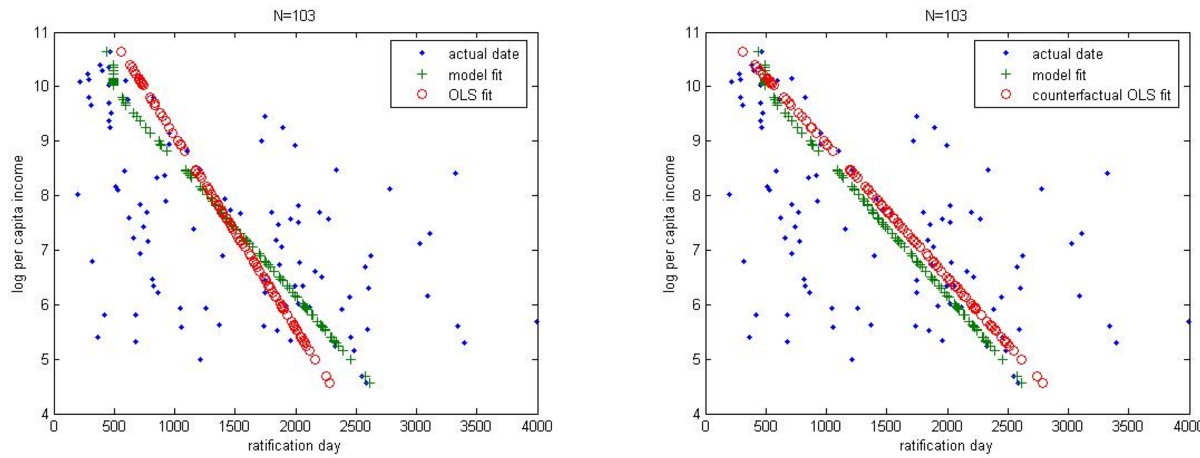
(a) (b)

Figure 3: Fitted Model

well. The model slightly overpredicts early ratification times and underpredicts ratification times in the top decile of the distribution.

6.3.3 Comparison to a non-strategic model

Figure 4 plots both the fitted strategic model and a linear prediction based on an OLS regression of the ratification date on all covariates included in the net cost term ϕ. For both models, predicted ratification times are plotted against log per capita income, holding all other covariates fixed at their mean values. The figure illustrates two points. First, the linear fit misses the clustering of ratification dates early on. Second, if the strategic model of ratification is correct, the non-strategic OLS estimates will be biased towards zero. This is because countries with a higher observable cost are higher in the order of ratification and thus receive a larger "subsidy" from other countries who have already joined. This subsidy is not accounted for in the non-strategic model, and hence induces a negative correlation between unobservable and observable parts of the cost. The bias is apparent in figure 4a where the linear fit implies a smaller coefficient on per capita income than the strategic model. To examine whether this difference is due to the different functional form assumptions underlying the models, I conduct the following experiment: Using the estimated parameters from the fitted model and drawing at random on the distribution of the structural unobservable ϵ, I simulate a non-strategic ratification history by setting $\gamma = 0$. The resulting OLS fit for this history, shown in figure 4b, implies virtually the same parameter as the strategic model. This suggests that it is attenuation bias, and not



(a) (b)

Figure 4: Strategic Model vs. Non-strategic Model

functional form, that is driving the differences in figure 4a.[25]

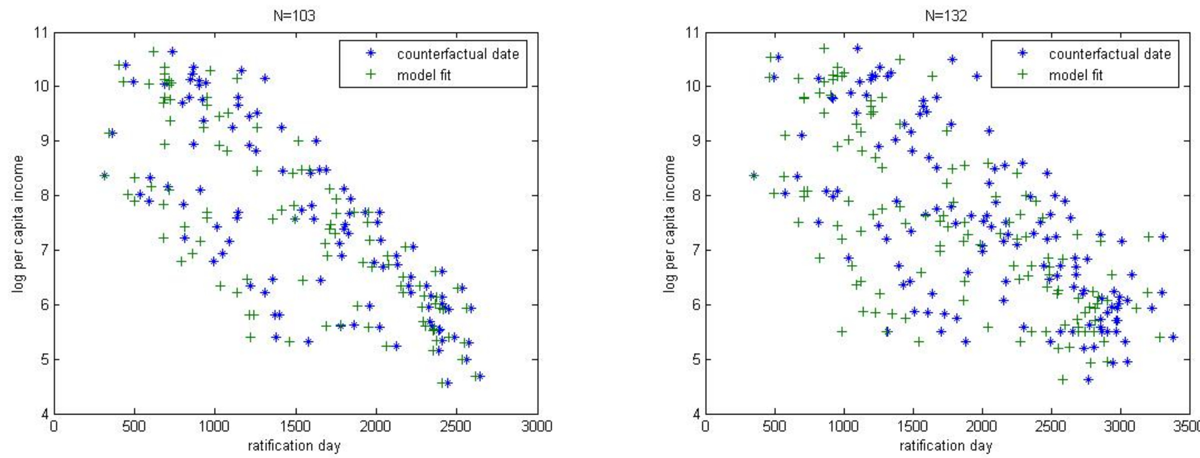
6.3.4 Counterfactual ratification times

How strong are the strategic complementarities implied by the estimates in table 4? To answer this question, I compare the fitted ratification times to counterfactual fitted ratification times obtained by setting $\gamma = 0$. Figure 5 plots both ratification histories against log per capita income. Strategic complementarities bring forward the ratification time and generate some clustering of ratification dates. The overall magnitude of this effect is moderate. In the smaller sample with 103 countries (Figure 5a) the mean ratification time decreases by 107 days (15 weeks) and the median by 115 days (16 weeks). This corresponds to about one eighth of the standard deviation of the time to ratification. The effect is larger in the more complete sample of 132 countries. Strategic complementarities reduce the mean of the time to ratification by 247 days (35 weeks) and the median by 309 days (44 weeks), which corresponds to one fifth and one forth of the standard deviation, respectively. The preferred estimate is 35 weeks, since it is based on a more representative sample.

6.3.5 Bounding the effect of equilibrium selection

Imposing strong renegotiation-proofness implicitly assumes perfect coordination on Pareto-improving outcomes at every stage of the game. To bound the effect of equilibrium selection

[25] Similarly, I verify that an estimate of $\hat{\gamma} > 0$ is not an artifact of the estimation routine and/or functional form assumptions by re-estimating the model using simulated non-strategic ratification histories. In each of these experiments, parameter γ is precisely estimated at its true value of 0.



(a) (b)

Figure 5: Counterfactual Ratification Times

on the magnitude of strategic complementarity, I re-estimate the model under the assumption of total coordination failure. In the context of the simple 3-period model shown in figure 1b, this means that players play (defect, defect, cooperate) rather than (defect, cooperate, cooperate). I obtain larger point estimates under this equilibrium assumption, in the large sample and 11.369 vs. 10.564 in the small sample and 16.768 vs. 9.225 in the larger sample. The difference is not statistically significant. In the large sample, the alternative equilibrium gives rise to a mean (median) reduction of the time to ratification by 55 (69) weeks. The preferred estimate of 35 weeks is hence viewed as a lower bound on the magnitude of strategic complementarities.

7 The Sources of Strategic Complementarities

The economics and international relations literatures have discussed specific government interaction effects that may create strategic complementarities. Some of these effects are strategic in nature whereas others are social reinforcement effects. This section employs the empirical framework developed so far to examine their role in determining the strength of strategic complementarities.

7.1 Economic Dependency and Trade Restrictions

Previous research has suggested that a country's ratification decision may be influenced by the decisions of countries that it depends upon economically. In the context of the early

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

$$w_{ij} = \frac{\text{exports from } i \text{ to } j}{\text{total exports from country } i} \tag{7.1}$$

using total commodity exports in 1986 taken from the World Trade Analyzer CD-Rom (StatCan, 1998).

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 (1997a) argues that this ban transformed trade in controlled substances into a club good whose benefits were exclusive to member states. Once the club reaches a critical size, trade with member states is gives higher benefits than trade with non-member states. If data on trade in controlled substances before 1987 were available, they could be used to gauge the effect of trade restrictions on participation. Unfortunately, the UN Ozone Secretariat does not make bilateral trade data available. This confines me to using the more aggregated data on exports of "Chemicals and Related Products" (also available from StatCan, 1998)) as a proxy variable.

The inclusion of trade data results in a slightly smaller data set. Descriptive statistics are reported in table 5 and show no significant differences to the larger data set. Table 6 reports the results obtained when estimating the model using both sets of trade-related weights. Columns 1 through 3 show the results when spillovers are calibrated on total commodity exports. The coefficient estimate for γ is virtually zero for all specifications. This result is consistent with earlier results by Beron et al. (2003), who found that bilateral exports had no significant effect on treaty membership. Moreover, the parameter estimate on LNCFC is negative and the estimated coefficients on LNPCY and LNPOP are sensitive to its inclusion, possibly due to collinearity of those variables.

Columns 4 through 6 of table 6 show a very similar pattern. Here, the w_{ij} are calibrated on bilateral exports of chemical products only. Again, the spillover parameter is virtually zero. One interpretation of this result is that trade restrictions had no effect on participation in the Montreal Protocol. However, since bilateral trade in chemical products is a fairly crude proxy for trade in substances controlled under the Montreal Protocol, it could also be that measurement error biases the coefficient estimate towards zero. Given the unavailability of pertinent trade data, the evidence on this question remains inconclusive.

7.2 Issue Linkage and Reputation Effects

The second transmission channel for endogenous effects considered here is the role of reputation and issue linkage. It is widely recognized that diplomats may choose to negotiate different topics jointly in order to achieve more stable outcomes (see Raiffa, 1982, Sebenius, 1983, Tollison and Willett, 1979). Cesar and de Zeeuw (1996) and Folmer et al. (1993) provide specific examples of how such interconnections can help to stabilize an international environmental agreement. The essence of this idea is that 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 get at this effect, I calibrate w_{ij} to the degree of involvement in R pre-existing international agreements,

$$w_{ij} = \frac{1}{R*(N-1)} \sum_{r=1}^{R} \mathbb{1}\{i \text{ and } j \text{ signed treaty } r\} \ln(GDP_j) \tag{7.2}$$

where logged GDP is included to control for the importance of a country. I focus on preexisting treaties in order to preclude potential simultaneity of the decision to join the Montreal Protocol and to ratify other treaties.[26]

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) 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" (p. 155). They examine participation in a self-enforcing international environmental agreement under the assumption that non-member states incur a non-environmental cost that is increasing in the number of signatories. In the present model, Hoel and Schneider's argument would predict that country i's reputation benefit from ratification is greater the more of its "peers" are among the signatories.[27] The weights (7.2) can thus be interpreted as a plausible (though not the only conceivable) calibration to estimate the magnitude of reputation effects discussed by Hoel and Schneider (1997). Ultimately, it appears challenging

[26] Specifically, I look at membership status in the following 11 international agreements: Charter of the United Nations; International Court of Justice; Law of Sea Convention; International Tribunal for the Law of Sea jurisdiction; International Convention on the Elimination of Racial Discrimination; Convention Against Torture and Other Cruel, Inhuman, or Degreading Treatment or Punishment; International Convention on Civil and Political Rights; Convention on the Political Rights of Women; Optional Protocol to the International Convention on Civil and Political Rights; Convention on the Prevention and Punishment of Genocide; Convention on the Elimination of all Forms of Discrimination against Women. This data is available online at http://untreaty.un.org. I thank Oona Hathaway for graciously providing me with her data set on human rights treaties.

[27] In a binary choice framework, a cost exclusive to non-members is equivalent to the benefit $\gamma \sum w_{ij}$ exclusive to member states.

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.

Table 7 reports the results obtained using the reputation weights given in eq. (7.2). All columns show that there is a positive and significant spillover effect whereas the estimated magnitude of this and some of the other parameters varies across specifications. The signs of the parameter estimates in column 1 correspond to those obtained in the previous tables for this specification. The inclusion of latitude leads to a positive coefficient on LNCFC which is consistent with the view that abatement cost is proportional to consumption. It also amplifies the magnitude of the coefficients on LNPCY, LNPOP, ART5 and the constant term. This may be due to the fact that most developed countries are located in higher latitudes, inducing a correlation with per capita income, high CFC consumption, and the dummy for developing countries, ART5. The sign on latitude itself is positive, which is at odds with the notion that higher latitude is associated with higher benefits. The collinearity issue adds to the fact that I need to estimate a complex nonlinear model using relatively few observations. Therefore, I exclude LATDEG in most specifications. Column 3 reports a more parsimonious specification where LNCFC is dropped. Compared to column 1, the estimated spillover coefficient doubles in size while most other coefficients are similar.

7.3 Fairness and Equity

Concerns about 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. In the so-called Berlin Mandate, the Conference of Parties to the Framework Convention on Climate Change agreed that developing countries should not be required to reduce their greenhouse gas emissions. While it must have been clear to most delegates that the exemption would result in globally inefficient provision of abatement, the majority subscribed to the view that it would not be fair to hold the developing world accountable for a problem caused by industrialized countries.

Fairness has received considerable attention in the recent economics literature (see,e.g., 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

$$w_{ij} = w_j = \frac{(\text{CFC emissions})_j}{\sum_{k=1}^{N} (\text{CFC emissions})_k}. \tag{7.3}$$

A positive coefficient γ means that accession of large emitters to the Montreal Protocol accelerates ratification by other countries.

Table 8 reports the evidence on the fairness hypothesis. Since the spillover matrix is calibrated on CFC consumption, I use the PRODUCER dummy in the cost term ϕ instead of LNCFC to mitigate possible multi-collinearity. The spillover term is positive and statistically significant in all three columns of the table. The estimates in column 1 are similar to the ones obtained with the treaty weighting matrix for this specification. The PRODUCER dummy is included in column 2. This decreases the size of the estimated spillover coefficient γ and increases that of most other coefficients. The coefficient estimate itself is not statistically significant.

Overall, the signs of the estimated coefficients on key determinants of net benefits – per capita income, population size, article 5 status – are robust across specifications and confirm intuition (see table 2). I find robust evidence of positive spillover effects that engender strategic complementarities in the ratification process. The model attributes spillovers to the frequency with which countries collaborate in other international agreements, and to their relative emission intensity, but not to the intensity of bilateral trade relations. These findings support the hypotheses that both reputation/issue linkage and equity may have played a role in countries' ratification decisions. By contrast, the results do not lend support to the conjecture that economic dependencies between countries or the treaty's trade restrictions were driving ratification.

8 Conclusions

This paper estimates the magnitude of reinforcement effects in the ratification of the Montreal Protocol on Substances that Deplete the Ozone Layer. In the presence of such effects, ratification is a strategic complement. A non-strategic analysis of the ratification process is inappropriate because it fails to account for forward-looking behavior on the part of governments. Therefore, I develop a strategic model of the timing of treaty ratification, which is amenable to structural estimation. The fitted model predicts that strategic complementarities reduced the average time to ratification by 35 weeks. Investigation of different sources of strategic complementarities demonstrates that ratification by country A is more likely to trigger ratification by country B if pre-existing international treaties have been ratified by both countries, or if country A has a large share in global emissions of ozone-depleting substances. The former points to concern about reputation; the latter shows a preference for equity. By contrast, there is no evidence that bilateral trade flows affected the strength of strategic complementarities.

Although there are a number of fundamental differences between climate change and stratospheric ozone depletion, the common features of these global environmental problems allow me to offer a few policy implications for the design of an effective global climate treaty. First, strategic complementarities may rein in the free-riding incentive inherent in the provision of abatement. If this effect is strong, fears that China and India will free-ride on a treaty with U.S. participation are unfounded. In contrast, ratification by the U.S. could break the political stalemate and trigger ratification by other countries. To be sure, this would require the U.S. to overcome its own incentive to free-ride. It appears that to date this incentive has been strong enough to outweigh the political cost this U.S. administration has been incurring since it withdrew from the Kyoto Protocol in 2001.

Second, from the outset of the international negotiations on climate change mitigation, equity has been playing an important role. The Framework Convention on Climate Change incorporates the principle of common but differentiated responsibility, recognizing the drastic imbalances across countries in both historical and current responsibilities for the build-up of the stock of greenhouse gases in the atmosphere. It seems very plausible that a Kyoto Protocol with U.S. participation will be perceived as more equitable by developing countries. My empirical results for the Montreal Protocol would thus suggest that ratification by the U.S. would increase the willingness on the part of developing countries to take on binding emission targets.

Finally, it has been suggested that trade sanctions be employed to punish non-compliance with a global climate treaty. The findings in this paper do not provide evidence of an association between trade relations and participation in the Montreal Protocol. With respect to total commodity trade, this is perhaps not surprising because the implicit threat of banning all imports from non-member states clashes with WTO rules and hence is not credible. However, since the UNEP Ozone Secretariat does not make pertinent data on trade in controlled substances available for research, I cannot rule out the possibility that the Montreal Protocol's ban on trade in those substances did enhance participation. It is conceivable that trade sanctions that are specific, narrowly-targeted, and compatible with the rules of the WTO are effective at enforcing compliance with a global climate treaty.

To conclude, this paper has developed a novel empirical framework for estimating strategic complementarities in the timing of public good provision. Since the underlying theoretical model is a discrete dynamic game with monotonically evolving payoffs, the framework is readily applicable to a host of questions in the social sciences that involve repeated interactions between a large number of agents in the presence of strategic complementarities. Examples include the adoption of new technologies that exhibit network effects, land-use conversion decisions that impose spatial externalities, and dynamic residential choice when peer effects matter. Furthermore, applying this framework to analyze strategic reinforcement effects in other international treaties is straightforward and an interesting topic for future research.

References

AFEAS: 2006, Production, sales, and atmospheric release of fluorocarbons through 2003, *Technical report*, Alternative Fluorocarbons Environmental Acceptability Study, Arlington, VA. Published online at `http://www.afeas.org/production_and_sales.html`.

Aguirregabiria, V. and Mira, P.: 2007, Sequential estimation of dynamic discrete games, *Econometrica* **75**(1), 1–53.

Auffhammer, M., Morzuch, B. and Stranlund, J.: 2005, Production of chlorofluorocarbons in anticipation of the Montreal Protocol, *Environmental and Resource Economics* **30**(4), 377–391.

Aumann, R. J.: 1959, Acceptable points in general cooperative n-person games, *in* R. D. Luce and A. W. Tucker (eds), *Contributions to the Theory of Games IV*, Vol. 40 of *Annals of Mathematics Studies*, Princeton University Press, Princeton, NJ, pp. 287–324.

Bajari, P., Benkard, C. L. and Levin, J.: 2007, Estimating dynamic models of imperfect competition, *Econometrica* **75**(5), 1331–1370.

Barrett, S.: 1994, Self-enforcing international environmental agreements, *Oxford Economic Papers* **46**(Special Issue), 878–94.

Barrett, S.: 1997a, Heterogeneous international environmental agreements, *in* C. Carraro (ed.), *International Environmental Negotiations. Strategic Policy Issues*, Edward Elgar, Cheltenham, UK, chapter 2, pp. 9–25.

Barrett, S.: 1997b, The strategy of trade sanctions in international environmental agreements, *Resource and Energy Economics* **19**(4), 345–361.

Barrett, S.: 2002, Consensus treaties, *Journal of Institutional and Theoretical Economics* **158**(4), 529–547.

Barrett, S.: 2003, *Environment and Statecraft*, Oxford University Press, New York.

Benedick, R. E.: 1998, *Ozone Diplomacy : New Directions in Safeguarding the Planet*, enlarged edn, Harvard University Press, Cambridge, MA.

Bergstrom, T., Blume, L. and Varian, H. R.: 1986, On the private provision of public goods, *Journal of Public Economics* **29**(1), 25–49.

Bernheim, B. D., Peleg, B. and Whinston, M. D.: 1987, Coalition-proof Nash equilibria I. Concepts, *Journal of Economic Theory* **42**(1), 1–12.

Bernheim, B. D. and Ray, D.: 1989, Collective dynamic consistency in repeated games, *Games and Economic Behaviour* **1**(4), 295–326.

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**(2), 286–97.

Berry, S.: 1992, Estimation of a model of entry in the airline industry, *Econometrica* **60**, 889–917.

Bolton, G. E. and Ockenfels, A.: 2000, Self-centered fairness in games with more than two players, *American Economic Review* **90**, 166–193.

Brock, W. A. and Durlauf, S. N.: 2001, Discrete choice models with interactions, *Review of Economic Studies* **68**, 235–260.

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**(3), 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*, Edward Elgar, Cheltenham, UK, chapter 7, pp. 158–173.

Chander, P. and Tulkens, H.: 1995, A core-theoretic solution for the design of cooperative agreements on transfrontier pollution, *International Tax and Public Finance* **2**(2), 279–294.

Collard-Wexler, A.: 2006, Demand fluctuations and plant turnover in the ready-mix concrete industry, *Working paper*, NYU, New York, NY.

Congleton, R. D.: 1992, Political institutions and pollution control, *Review of Economics and Statistics* **74**, 412–21.

Cooper, R. and John, A.: 1988, Coordinating coordination failures in Keynesian models, *Quarterly Journal of Economics* **103**, 441–461.

Crutzen, P. J.: 1970, The influence of nitrogen oxides on the atmospheric ozone content, *Quarterly Journal of the Royal Meteorological Society* **96**, 320–25.

De Paula, A.: 2006, Inference in a synchronization game with social interactions, *Working paper*, University of Pennsylvania, Philadelphia, PA.

Demsetz, H.: 1967, Toward a theory of property rights, *American Economic Review* **57**, 347–359.

Einav, L.: 2003, Not all rivals look alike: Estimating an equilibrium model of the release date timing game, *Working paper*, Stanford University, Palo Alto, CA.

Farrell, J. and Maskin, E.: 1989a, Renegotiation in repeated games, *Games and Economic Behaviour* **1**(4), 327–360.

Farrell, J. and Maskin, E.: 1989b, Renegotiation-proof equilibrium: Reply, *Journal of Economic Theory* **49**, 376–378.

Fehr, E. and Schmidt, K. M.: 1999, A theory of fairness, competition and cooperation, *Quarterly Journal of Economics* **114**(3), 817–868.

Finus, M.: 2001, *Game Theory and International Environmental Cooperation*, Edward Elgar, Cheltenham, UK.

Folmer, H., v. Mouche, P. and Ragland, S.: 1993, Interconnected games and international environmental problems, *Environmental and Resource Economics* **3**(4), 313–335.

Fudenberg, D., Gilbert, R., Stiglitz, J. and Tirole, J.: 1983, Preemption, leapfrogging, and competition in patent races, *European Economic Review* **22**, 3–31.

Fudenberg, D. and Maskin, E.: 1986, The folk theorem in repeated games with discounting or with incomplete information, *Econometrica* **54**(3), 533–554.

Fudenberg, D. and Tirole, J.: 1985, Preemption and rent equalization in the adoption of new technology, *Review of Economic Studies* **52**, 383–402.

Ghemawat, P. and Nalebuff, B.: 1985, Exit, *RAND Journal of Economics* **16**(2), 184–94.

Glaeser, E. and Scheinkman, J.: 2002, Non-market interactions, *Working paper*, Department of Economics, Harvard University, Cambridge, MA.

Goffe, W., Ferrier, G. and Rogers, J.: 1994, Global optimization of statistical functions with simulated annealing, *Journal of Econometrics* **60**(1/2), 69–99.

Hammitt, J. K.: 2000, Are the costs of proposed environmental regulations overestimated? Evidence from the CFC phaseout, *Environmental and Resource Economics* **16**, 281–301.

Hardin, G.: 1968, The tragedy of the commons, *Science* **162**, 1243–1248.

Hathaway, O. A.: 2004, Do environmental treaties make a difference?, *Working paper*, Yale Law School, New Haven, CT.

Heal, G. and Kunreuther, H.: 2007, Social reinforcement: Cascades, entrapment and tipping, *Working Paper 13579*, NBER, Cambridge, MA.

Hendricks, K., Weiss, A. and Wilson, C.: 1988, The war of attrition in continuous time with complete information, *International Economic Review* **29**(4), 663–680.

Hoel, M.: 1992, International environmental conventions: The case of uniform reductions of emissions, *Environmental and Resource Economics* **2**(3), 141–159.

Hoel, M. and Schneider, K.: 1997, Incentives to participate in an international environmental agreement, *Environmental and Resource Economics* **9**(2), 153–70.

Jia, P.: 2005, What happens when Walmart comes to town: An empirical analysis of the discount retailing industry, *Job market paper*, Yale University, New Haven, CT.

Kolstad, C. D.: 2007, Systematic uncertainty in self-enforcing international environmental agreements, *Journal of Environmental Economics and Management* **53**(1), 68–79.

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.

Lovelock, J. E., Maggs, R. J. and Wade, R. J.: 1973, Halogenated hydrocarbons in an over the Atlantic, *Nature* **245**, 194.

Mazzeo, M.: 2002, Product choice and oligopoly market structure, *RAND Journal of Economics* **33**(2), 1–22.

McFadden, D.: 1989, A method of simulated moments for estimation of discrete response models without numerical integration, *Econometrica* **57**(5), 995–1026.

McGinty, M.: 2007, International environmental agreements among asymmetric nations, *Oxford Economic Papers* **59**(1), 45–62.

Milgrom, P. and Roberts, J.: 1990, Rationalizability, learning, and equilibrium in games with strategic complementarities, *Econometrica* **58**(6), 1255–1277.

Molina, M. and Rowland, F. S.: 1974, Stratospheric sink for chlorofluoromethanes: Chlorine atom-catalyzed destruction of ozone, *Nature* **249**, 810–12.

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**(3), 331–349.

Murdoch, J. C., Sandler, T. and Sargent, K.: 1997, A tale of two collectives: Sulphur versus nitrogen oxides emission reduction in Europe, *Economica* **64**, 281–301.

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.

Pakes, A., Ostrovsky, M. and Berry, S. T.: 2007, Simple estimators for the parameters of discrete dynamic games, with entry/exit examples, *RAND Journal of Economics* **38**(2), 373–399.

Pakes, A. and Pollard, D.: 1989, Simulation and the asymptotics of optimization estimators, *Econometrica* **57**(5), 1027–1057.

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

Pearce, D.: 1987, Renegotiation-proof equilibria: Collective rationality and interemporal cooperation, *mimeo*, Yale University, New Haven, CT.

Pearce, D. G.: 1990, Repeated games: Cooperation and rationality, *in* J.-J. Laffont (ed.), *Advances in Economic Theory, Sixth World Congress*, Vol. 1, Cambridge University Press, Cambridge, UK, chapter 4, pp. 132–174.

Pesendorfer, M. and Schmidt-Dengler, P.: 2006, Asymptotic least squares estimators for dynamic games, *Working paper*, London School of Economics, London, UK.

Raiffa, H.: 1982, *The Art and Science of Negotiation*, Harvard University Press, Cambridge, MA.

Reinganum, J.: 1981a, Market structure and the diffusion of new technology, *Bell Journal of Economics* **12**, 618–624.

Reinganum, J.: 1981b, On the diffusion of a new technology: A game-theoretic approach, *Review of Economic Studies* **48**, 295–405.

Rubio, S. J. and Ulph, A.: 2006, Self-enforcing international environmental agreements revisited, *Oxford Economic Papers* **58**, 233–263.

Ryan, S. P.: 2006, The costs of environmental regulation in a concentrated industry, *Working paper*, MIT, Cambridge, MA.

Samuelson, P. A.: 1954, The pure theory of public expenditure, *The Review of Economics and Statistics* **36**(4), 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**(2), 281–316.

Smith, J. M.: 1974, The theory of games and the evolution of animal conflicts, *Journal of Theoretical Biology* **47**, 209–221.

StatCan: 1998, World trade analyzer, CD-Rom.

Stolarski, R. S. and Cicerone, R. J.: 1974, Stratospheric chlorine: A possible sink for ozone, *Canadian Journal of Chemistry* **52**, 1610–1615.

Sweeting, A.: 2005, Coordination games, multiple equilibria and the timing of radio commercials, *Working paper*, Northwestern University, Chicago, IL.

Sweeting, A.: 2006, Coordination, differentiation and the timing of radio commercials, *Journal of Economics and Management Strategy* **15**(4), 909–942.

Tollison, R. D. and Willett, T. D.: 1979, An economic theory of mutually advantageous issue linkage in international negotiations, *International Organization* **33**(4), 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 University Press, Princeton, NJ.

UNEP: 2004, United Nations Environment Programme GEO Data Portal, available online. http://geodata.grid.unep.ch/.

van Damme, E.: 1989, Renegotiation-proof equilibria in repeated prisoners' dilemma, *Journal of Economic Theory* **47**(1), 206–217.

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**(3), 377–411.

Wagner, U. J.: 2007, The voluntary provision of a pure public good? Another look at CFC emissions and the Montreal Protocol, *Oxford Economic Papers* **forthcoming**.

Zhou, L.: 1994, The set of Nash equilibria of a supermodular game is a complete lattice, *Games and Economic Behavior* **7**, 295–300.

Appendix A: Proofs

Proof of Theorem 1 (Existence of a pure-strategy NE in the stage game). Notice first that 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. Under both assumptions 3 and 4, 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, quoted in appendix B) and hence $\Gamma(t)$ is a supermodular game $\forall t \in T$. Notice further that the payoff function $\pi_i(a_i, a_{-i}, t)$ is upper semi-continuous in a_i on A_i for each $a{-}i \in A_{-i}$ and each $i \in I$. Under these conditions, theorem 4.2.1 from Topkis (1998, see appendix B) – which summarizes results by Topkis (1979) and Zhou (1994) – can be invoked to establish that the set of Nash equilibria is a complete lattice that contains a greatest and a least element. ∎

Proof of theorem 2. In a dominant strategy equilibrium, the best response correspondence $R(\cdot, t)$ 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 3 we have that

$$\pi_i(R_i(a'_{-i}, t), a'_{-i}, t) - \pi_i(R_i(a_{-i}, t), a'_{-i}, t) = \pi_i(R_i(a'_{-i}, t), a_{-i}, t) - \pi_i(R_i(a_{-i}, t), a_{-i}, t). \tag{8.1}$$

By the definition of the best response function (3.3), 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 assumption 7 rules out ties for different actions, eq. (8.1) 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.

The decision rule picks the dominant strategy for each player because $a_i^*(t) = R_i(a_{-i}^*, t) = R_i(\mathbf{0}, t)$, where the first equality follows from the definition of Nash equilibrium and the second one from the fact that the best response function is a constant. Suppose now that the dominant action is decreasing with t. Then there exist two periods $t', t'' \in T$ with $t' < t''$ and some player $i \in I$ such that $\pi_i(1, 0, t') \geq \pi_i(0, 0, t')$ and $\pi_i(1, 0, t'') < \pi(0, 0, t'')$. This implies

$$\pi_i(1, 0, t') - \pi_i(1, 0, t'') \geq \pi_i(0, 0, t') - \pi(0, 0, t'')$$

which contradicts assumption 1. It follows that the dominant action is weakly increasing in t for all players, and hence the Nash equilibrium action vector a^* is weakly increasing in t. ∎

Proof of theorem 3. From the definition of t_i^0 it follows that full cooperation is the unique NE of the stage game in period t_N^0 and, by the monotonicity property stated in theorem 2, in all subsequent periods. By definition, indefinite repetition of this stage-game equilibrium constitutes a subgame-perfect continuation equilibrium at t_N^0.

This SPNE is WRP because it is not dominated by any of its continuation equilibria. The equilibrium is also SRP since it stipulates an action profile $a^N = \mathbf{1}$ that strictly dominates every other profile in the stage game, and hence is undominated by any other WRP equilibrium. To see this, consider a profile $a' \in A; a' \leq a^N, a' \neq a^N$. By assumption 2, players who take the same action under both profiles strictly prefer a^N. Moreover, players who do not cooperate under a' are worse off: $\pi_j(1, a_{-j}^N, t) > \pi_j(0, a_{-j}^N, t) > \pi_j(0, a'_{-i}, t)$ where the first inequality follows from the fact that cooperation is the dominant action and the second inequality follows from the assumption of positive spillovers. Since this is true for any such profile a', full cooperation is the unique SRP equilibrium of the continuation game beginning in period t_N^0, $G(t_N^0)$.

This feature likens the solution of the overall game to that of a finitely repeated game. If equilibrium strategies are SRP, players know that everyone joins the coalition once play reaches period t_N^0. Using a simple backward induction argument, it is straightforward to see that the unique NE of the stage game must be played at each of the decision nodes preceding t_N^0. Hence, in the unique SPNE that satisfies strong renegotiation-proofness, each player i joins the coalition in period t_i^0, defined in eq. (3.4). ∎

Proof of theorem 4. The proof is comprised of three steps. First, it is shown that there exists a time $t_N^0 < \infty$ such that the continuation game has a unique SRP equilibrium in which every player cooperates. The overall game can hence be solved from period t_N^0 backwards, just like a finite horizon game. In particular, perfection requires that a Nash equilibrium be played in every stage game before t_N^0. Next, I prove that strong renegotiation-proofness requires that the greatest Nash equilibrium be played. Finally, 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: All players cooperate in finite time Analogous to the proof of theorem 3 it can be shown that full cooperation is the unique SRP equilibrium of the continuation game $G(t_N^0)$. This implies that the only credible "punishment" that the coalition can inflict on a player who deviates in period $t_N^0 - 1$ is to play $a^N = \mathbf{1}$ indefinitely; in other words, the deviation goes unpunished. Accordingly, the only action profile that can be sustained at node $t_N^0 - 1$ in SPNE is a Nash equilibrium of the stage game $\Gamma(t_N^0 - 1)$. Backward iteration of this argument establishes that a stage-game Nash equilibrium must be played in all previous periods, too. In addition, the refinement of strong renegotiation-proofness requires that none of these stage-game equilibria be Pareto-dominated by another stage-game equilibrium, for if there were a WRP profile $\tilde{s}$ inducing a dominated NE in the stage game at t', the profile $\tilde{s}'$ that induces the dominant NE in stage t' and is otherwise identical to $\tilde{s}$ would be WRP, too, and would give a strictly higher payoff than $\tilde{s}$. Hence, $\tilde{s}$

cannot be SRP.

Step 2: The largest Nash equilibrium of the stage game Pareto-dominates all other Nash equilibra The set of NE of the game $\Gamma(t)$ is a complete lattice with a greatest and a least element (Topkis, 1979, Zhou, 1994). Suppose now that $a^*, a^{**} \in A$ are both NE of the stage game and $a^* \geq a^{**}$. To see that a^* strictly dominates a^{**}, notice first that, $\forall i \in I$ such that $a_i^{**} = a_i^*$, assumption 2 of positive spillovers implies that $\pi_i(a^*, t) > \pi_i(a^{**}, t)$. For all other players $j \in I$ such that $a_j^* > a_j^{**}$ consider the following two cases.

1. $a_{-j}^* \geq a_{-j}^{**}$: $\pi_j(a_j^*, a_{-j}^*, t) \geq \pi_j(a_j^{**}, a_{-j}^*, t) > \pi_j(a_j^{**}, a_{-j}^{**}, t)$ where the first inequality follows from the definition of NE and the second inequality from assumption 2.

2. $a_{-j}^* = a_{-j}^{**}$: $\pi_j(a_j^*, a_{-j}^*, t) > \pi_j(a_j^{**}, a_{-j}^*, t) = \pi_j(a_j^{**}, a_{-j}^{**}, t)$ where the first inequality holds by assumption 7.

It follows that the largest NE gives each player a strictly higher payoff than any other NE.

Step 3: The profile s^* induces the largest Nash equilibrium in every stage game Finally, it needs to be shown that the strategy profile s^* induces the largest NE in every period. Theorem 4.2.2 in Topkis (1998, given in appendix B) establishes that the largest NE of a supermodular game is increasing in the parameter t provided that assumption 1 holds. Specifically, monotonicity implies that a player who cooperates in the largest NE in some period t' will not revert her decision in the largest NE in any later period $t'' > t'$. Therefore, solving for the equilibrium path of G boils down to finding, for each player $i \in N$, the period in which i cooperates for the first time in the largest Nash equilibrium.

To begin, consider two useful implications of Topkis' theorem. First, since full cooperation is a Nash equilibrium at $t_{(N)}^*$, the theorem says that full cooperation – the greatest element of the action space A – is the largest Nash equilibrium in every subsequent period. Second, eq. (3.8) implies that in period $t_{(N)}^* - 1$ player $j_{(N)}$'s best response to all other players cooperating is to defect. The monotonicity theorem then tells us that $j_{(N)}$ cannot be part of the largest (and hence: any) Nash equilibrium in any earlier period $t < t_{(N)}^* - 1$.

Next, the principle of induction is used to show that, for each $k = 1, \ldots, N-1$ the strategy profile s^* induces the largest NE of the stage game in periods $t_{(N-k)}^*$ through $\max[t_{(N-k)}^*, t_{(N-k+1)}^* - 1]$. For 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. As shown above, this is the largest NE of the stage game from period $t_{(N)}^*$ onwards. Suppose now that $t_{(N-1)} < t_{(N)}^*$. Then $t_{(N-1)}^* = t_{(N-1)}$ and, from eq. (3.5), $\pi_i(K_{N-1} \backslash \{i\}, t_{(N-1)}) \leq \pi_i(K_{N-1}, t_{(N-1)})$ $\forall i \in K_{N-1}$. Therefore, none of the players in K_{N-1} has an incentive to deviate. Moreover, from assumption 1 and eq. (3.8) it follows that $\forall t < t_{(N)}^*$,

$$\Delta\pi_{(N)}(K_N \backslash \{j_{(N)}\}, t) \leq \Delta\pi_{(N)}(K_N \backslash \{j_{(N)}\}, t_{(N)}^* - 1) < 0.$$

Hence, cooperation by the set of players K_{N-1} constitutes a NE of the stage game in $t_{(N-1)}$. Moreover, since it contains all players except $j_{(N)}$ it is also the largest NE in period $t_{(N-1)}$, and monotonicity implies that this is true for all $t_{(N-1)} \leq t < t_{(N)}^*$. From eq. (3.8), notice that player $j_{(N-1)}$ does not have an incentive to cooperate in any period before $t_{(N-1)}$.

Assume now that, for some k, $1 \leq k \leq N-1$ the strategy profile s^* induces the largest Nash equilibrium in the stage game from period $t_{(N-k)}^*$ through $\max[t_{(N-k)}^*, t_{(N-k+1)}^*]$. The inductive step is to prove that, given this hypothesis, the profile s^* induces the largest NE in each of the periods $t_{(N-(k+1))}^*$ through $\max[t_{(N-(k+1))}^*, t_{(N-k)}^* - 1]$.

Suppose first that $t_{(N-(k+1))} \geq t_{(N-k)}^*$. The claim that the strategy profile s^* induces the largest NE of the stage game in period $t_{(N-(k+1))}^* = t_{(N-k)}^*$ then follows directly 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 coalition K_{N-k-1}. From eq. (3.5),

$$\pi_i(K_{N-k-1} \backslash \{i\}, t_{(N-k-1)}) \leq \pi_i(K_{N-k-1}, t_{(N-k-1)}) \quad \forall i \in K_{N-k-1},$$

i.e. none of the players in K_{N-k-1} has an incentive to defect. To see that no additional player wishes to join the coalition, notice that $\forall l$, $0 \leq l \leq k$,

$$\begin{aligned}
&\pi_{(N-l)}(K_{N-k-1} \cup \{j_{(N-l)}\}, t_{(N-k-1)}) - \pi_{(N-l)}(K_{N-k-1}, t_{(N-k-1)}) \\
\leq &\pi_{(N-l)}(K_{N-l}, t_{(N-k-1)}) - \pi_{(N-l)}(K_{N-l-1}, t_{(N-k-1)}) \\
\leq &\pi_{(N-l)}(K_{N-l}, t^*_{(N-l)} - 1) - \pi_{(N-l)}(K_{N-l-1}, t^*_{(N-l)} - 1) \\
\leq &\pi_{(N-l)}(K_{N-l}, t_{(N-l)} - 1) - \pi_{(N-l)}(K_{N-l-1}, t_{(N-l)} - 1) < 0
\end{aligned} \tag{8.2}$$

where the first inequality follows from the assumption of increasing differences 4 and the fact that $K_{N-k-1} \subseteq K_{N-l}$, the second and third inequalities follow from assumption 1 and the fact that $t_{(N-k-1)} < t^*_{(N-l)} \leq t_{(N-l)}$, and the fourth inequality follows from eq. (3.8). Thus, the set of cooperating players K_{N-k-1} induced by s^* constitutes a NE of the stage game in period $t_{(N-k-1)}$. Since none of the larger coalitions K_{N-l}, $0 \leq l \leq k$ is a NE at stage $t_{(N-k-1)}$, it follows that K_{N-k-1} is induced by the greatest NE in period $t_{(N-k-1)}$. Finally, monotonicity implies that this is true for all $t_{(N-k-1)} \leq t < t^*_{(N-k)}$.

By the principle of induction, the claim must be true for all $k = 1, \ldots, N-1$. To determine the greatest Nash equilibrium in periods $t = 0, \ldots, t^*_{(1)} - 1$, recall that, from eq. (3.8), player $j_{(1)}$ has no incentive to join the coalition at any time before $t^*_{(1)}$. Therefore, in the greatest NE equilibrium before $t^*_{(1)}$ no player cooperates. This concludes the proof that s^* induces the largest Nash equilibrium in every stage game. From step 2, s^* thus induces the unique Pareto-dominant Nash equilibrium in every stage and, by the argument put forth in step 1, this profile constitutes the unique SRP equilibrium of the game. ■

Proof of theorem 5. By definition, every CPNE is also a NE, but the converse is not necessarily true. For all $t \in T$, the set of NE of the stage game $\Gamma(t)$ is a complete lattice with a greatest and a least element (Topkis, 1979, Zhou, 1994). It was shown above that the largest NE gives each player a strictly higher payoff than any other NE, hence only the former can be a CPNE.

It remains to be shown that the largest NE is self-enforcing, i.e. that there exists no profitable deviation by any proper subset of players $K \subset I$ that is itself robust to further deviations. Denote by K^* the set of cooperating players induced by a^*. Suppose first that a subset $C \subseteq K^*, |C| \geq 2$ of players change their actions from 1 to 0. The deviation is not self-enforcing since it gives every $i \in C$ a strictly lower payoff

$$\pi_i(K^* \backslash C, t) < \pi_i(K^* \backslash \{i\}, t) \leq \pi_i(K^*, t)$$

where the first inequality follows from assumption 2 and the second inequality from the definition of NE.

Next, consider a deviation by a subset of players $D \subseteq I \backslash K^*$ who change their action from 0 to 1. A necessary condition for the deviation to be self-enforcing is that no player in D must have an incentive to revert back to 0 unilaterally, i.e.

$$\pi_i(K^* \cup D, t) \geq \pi_i(K^* \cup D \backslash \{i\}, t) \quad \forall i \in D.$$

Moreover, none of the players in K^* wants to withdraw upon accession of D since

$$0 < \pi_i(K^*, t) - \pi_i(K^* \backslash \{i\}, t) \leq \pi_i(K^* \cup D, t) - \pi_i((K^* \cup D) \backslash \{i\}, t) \quad \forall i \in K^*,$$

where the first inequality follows from the definition of NE and assumption 7 and the second inequality from the assumption of increasing differences 4. The best response of the remaining players $I \backslash (K^* \cup D)$ may or may not be to join the larger coalition, depending on the strength of the complementarities. In either case, this does not change the optimal action by all players in K^* and D, since best response functions are weakly increasing and these players already choose their high action. Consequently, if the deviation by D is self-enforcing, there exists another Nash equilibrium $a^{**} \geq a^*$, which contradicts the fact that a^* is the largest Nash equilibrium of the stage game. Therefore, the deviation by D cannot be self-enforcing.

A final deviation that needs to be considered is a simultaneous deviation by a subset of players $C \subseteq K^*$ who withdraw from the coalition and a subset of players $D \subseteq I \backslash K^*$ who join the coalition. Clearly, the deviation is vulnerable to the players in C reverting back to cooperating since doing so gives each of them a strictly higher payoff:

$$\begin{aligned}\pi_i(K^* \cup D, t) - \pi_i(K^* \backslash C \cup D, t) &> \pi_i(K^* \cup D, t) - \pi_i\left((K^* \cup D) \backslash \{i\}, t\right) \\ &\geq \pi_i(K^*, t) - \pi_i(K^* \backslash \{i\}, t) \\ &\geq 0 \qquad \forall i \in C \subseteq K^*. \end{aligned} \tag{8.3}$$

The first and second inequalities follow from assumptions 2 and 4 (or 3), respectively. The third inequality follows from the definition of NE. The reversion by C is a self-enforcing deviation since any subsequent deviation by proper subsets of C gives each member a strictly lower payoff. This is true because the inequality (8.3) holds for all players in any subset C' of K^* and, hence, for all $C' \subset C$. This establishes that the largest NE of the stage game is the unique CPNE of the stage game. ■

Proof of theorem 6. The proof is by showing that if $a^o \leq a^*$ or a^o, a^* are unordered then a^o is not a maximizer of $W(a, t)$ in A. Denote by K^* the set of players who cooperate in a^*, K^o the set of players who cooperate in the optimum, $C = (K^* \cup K^o) \backslash K^*$ the set of players who begin to cooperate and by $D = (K^* \cup K^o) \backslash K^o$ the set of players in K^* who revert to defection. Clearly, $a^o \geq a^*$ is equivalent to $D = \emptyset$. Suppose now that $|D| > 0$. Then the change in the aggregate payoff is given by:

$$\Delta W = \underbrace{[W(K^o, t) - W(K^* \cup C, t)]}_{\Delta_2} + \underbrace{[W(K^* \cup C, t) - W(K^*, t)]}_{\Delta_1} \geq 0 \tag{8.4}$$

where

$$\Delta_1 = \sum_{j \in C} \underbrace{[\pi_j(K^* \cup C, t) - \pi_j(K^*, t)]}_{\gtreqless 0} + \sum_{j \in I \backslash C} \underbrace{[\pi_j(K^* \cup C, t) - \pi_j(K^*, t)]}_{\geq 0} \tag{8.5}$$

$$\Delta_2 = \sum_{j \in D} \underbrace{[\pi_j(K^o, t) - \pi_j(K^* \cup C, t)]}_{\leq 0} + \sum_{j \in I \backslash D} \underbrace{[\pi_j(K^o, t) - \pi_j(K^* \cup C, t)]}_{<0}. \tag{8.6}$$

The first term in (8.5) is ambiguous whereas the second term is nonnegative and strictly positive if $C \neq \emptyset$ due to assumption 2. The first term in (8.6) is nonpositive since Nash equilibrium implies that cooperation is individually rational given K^* for each of the players in $D \subseteq K^*$ and withdrawing from the even larger coalition $K^* \cup C$ cannot make any player $j \in D$ better off under assumptions 3 or 4. For all other players, the assumptions of positive spillovers 2 and $|D| > 0$ imply that the second term in (8.6) is negative and hence $\Delta_2 < 0$. Now consider the change in aggregate payoffs when going from the profile $a' \equiv a^* \vee a^o$ (which is in A by definition of a complete lattice) to profile a^o:

$$W(a^o, t) - W(a', t) = W(K^o, t) - W(K^* \cup C) = \Delta_2 < 0$$

The decrease in aggregate payoff contradicts the assumption that a^o is a maximizer of $W(a, t)$ in A. ■

Proof of theorem 7 (Limit of the discrete-time game). Suppose that the SRP equilibrium provision 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 provision time of player i in the game ${}^kG$ can be obtained by simply relabeling 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 $N_\epsilon = \lceil -\frac{\ln \epsilon}{\ln 2} \rceil$. By the same token, $x_{k'} - \tilde{t}_i^* < 2^{-k'} < 2^{-k}$ for all $k' > k$. ■

Proof of theorem 8. Consider a pair of heterogenous players i and j with (limit) equilibrium provision times $\tilde{t}_i^*, \tilde{t}_j^*$, respectively. From theorem 7, provision 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 $\tilde{t}_i^* = \tilde{t}_j^*$. Recall that, in the absence of strategic complementarities (assumption 3), the relative payoff to cooperation is constant in other players' actions: $\forall a_{-i} \in A_{-i}, i \in I$; $r(\phi_i, a_{-i}, \tilde{t}) = r(\phi_i, \tilde{t})$. The equilibrium conditions 3.15, $r(\phi_l, \tilde{t}_l^0) = 0 \quad \forall l \in I$ can thus be solved for $\tilde{t}_l^0 = r^{-1}(0, \phi_l)$ where assumptions 1 and 8 have been invoked to invert r. Therefore, $\tilde{t}_i^0 = \tilde{t}_j^0$ is equivalent to $r^{-1}(0, \phi_i) = r^{-1}(0, \phi_j)$. Since ϕ is a continuous random variable and r is strictly monotonic in ϕ, this event has probability zero. In contrast, if payoff functions exhibit increasing differences (assumption 4) then it follows from the recursive definition of $\tilde{t}^*_{(m)}$ in eq. (3.14) that limit provision times can be identical even among asymmetric players if strategic complementarities are sufficiently strong. ■

Proof of theorem 9. For $\gamma > 0$, the proof is by showing that the payoff function satisfies all assumptions required for applying theorem 4. First, notice that $\Delta\pi_i(\cdot) = -\phi_i + e^{\lambda t} + \gamma \sum_{j \neq i} w_{ij} a_j$. Assumption 1 is satisfied since $\lambda e^{\lambda t} > 0$ and $\lambda^2 e^{\lambda t} > 0$. Consider any two action profiles (a_i, a_{-i}), $(a_i, a'_{-i}) \in A$ such that $a_{-i} \leq a'_{-i}$. Assumption 2 holds since $EB_i(a_{-i}(t), t)$ is strictly increasing in a_{-i}. Assumption 4 holds because

$$\Delta\pi_i(a_{-i}, t) \leq -\phi_i + e^{\lambda t} + \gamma \sum_{j \neq i} w_{ij} a_j \leq -\phi_i + e^{\lambda t} + \gamma \sum_{j \neq i} w_{ij} a'_j \leq \Delta\pi_i(a'_{-i}, t).$$

Assumption 7 holds with probability one, since the probability that $-\phi_i + e^{\lambda t} + \gamma \sum_{j \neq i} w_{ij} a'_j$ for some $t \in T$ is zero for all $i \in I$, given the distributional assumptions on the vector ϵ. Notice that $\pi_i(1, \mathbf{0}, 0) \leq -\exp(x_i'\beta + \epsilon) < 0 \leq \pi_i(0, 0, 0)\ \forall i \in I$, which guarantees that assumption 5 is satisfied. Finally, it is assumed that players are sufficiently impatient $r_i > \lambda$, $\forall i \in I$ so that the boundedness condition 6 holds. The claim follows from theorem 4.

If $\gamma = 0$, then assumption 3 instead of 4 is satisfied. From theorem 3 and eq. (3.4), the equilibrium ratification time for country i is given by the smallest integer to satisfy $0 \leq -\phi_i + e^{\lambda t}$, given by $t_i^0 = \lceil \frac{1}{\lambda}(\ln \phi_i) \rceil$. This coincides with the t_i^* defined in theorem 9. ■

Appendix B: Useful Results on Supermodular Games

Supermodular functions on a chain *(Topkis, 1998, Corollary 2.6.1, p. 45)*

If X_i is a chain for $i = 1, \ldots, n$ and $f(x)$ has (strictly) increasing differences on $\times_{i=1}^n X_i$, then $f(x)$ is (strictly) supermodular on $\times_{i=1}^n X_i$.

Nash equilibrium *(Topkis, 1998, Theorem 4.2.1, p. 181)*

If $(N, S, \{f_i : i \in N\})$ is a supermodular game, the set S of feasible joint strategies is nonempty and compact, and the payoff function $f_i(y_i, x_{-i})$ is upper semicontinuous in y_i on $S_i(x_{-i})$ for each x_{-i} in S_{-i} and each i, then the set of equilibrium points is a nonempty complete lattice and a greatest and a least equilibrium point exist.

Monotone comparative statics *(Topkis, 1998, Theorem 4.2.2, p. 183)*

Suppose that T is a partially ordered set and $(N, S, \{f_i^t : i \in N\})$ is a collection of supermodular games parameterized by t in T where in game t the payoff function for each player i is $f_i^t(x)$ and the set of feasible joint strategies is S^t. The set S^t of feasible joint strategies is nonempty and compact for each t in T and is increasing in t on T. Let S^t_{-i} and $S_i^t(x_{-i})$ denote the dependence of S_{-i} and $S_i(x_{-i})$ on the parameter t. For each player i and each x_{-i} in S^t_{-i}, the payoff function $f_i^t(y_i, x_{-i})$ is upper semicontinuous in y_i on $S_i^t(x_{-i})$ for each t in T and has increasing differences in (y_i, t) on $(\cup_{t \in T} S_i^t) \times T$. Then there exists a greatest equilibrium point and a least equilibrium point for each game t in T, and the greatest (least) equilibrium point for game t is increasing in t on T.

Appendix C: Tables

Table 1: Average annual percentage growth rates of worldwide CFC production

	time period		
substance	1931-1977	1978-1986	1986-2003
CFC-11	20.63	1.0	-27.7
CFC-12	14.25	0.4	-20.4
CFC-113	n.a.	10.7	-34.1
CFC-114	n.a.	4.0	-23.8
CFC-115	n.a.	3.9	-33.9

source: AFEAS (2006) and own calculations

Table 2: Expected signs for the parameters of the net cost of complyingϕ

Variable	Description	Effect on net cost ϕ	Rationale
LNPCY	log per capita income	–	shifts demand for environmental quality
LNPOP	log population	–	shifts total benefits from abatement
LNCFC	log CFC consumption	+	abatement targets are proportional to consumption
PRODUCER	producer country	+/–	declining market for CFCs / profitable substitutes
ART5	Article 5 country	–	10-year grace period to meet abatement targets
LONDON_ART5	Article 5 country post London	+/–	n/a
LATDEG	latitude in degrees	–	stronger ozone depletion in higher latitudes

Table 3: Summary statistics

	A. $N = 103$				B. $N = 132$			
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
DAY	1449.825	906.145	197	3992	1753.242	1269.76	197	5536
LNPCY	7.506	1.633	4.573	10.650	7.456	1.598	4.621	10.708
LNPOP	2.095	1.617	-1.415	6.972	2.100	1.592	-2.349	7.020
LNCFC	5.869	2.718	0.833	12.631				
PRODUCER	0.165	0.373	0	1				
ART5	0.709	0.457	0	1	0.689	0.465	0	1
LONDON_ART5	0.485	0.502	0	1	0.508	0.502	0	1
LATITUDE	24.913	17.852	0	90				

$N = 103$. Clustering (day): Italy, Netherlands, Denmark, Spain (457); France, Switzerland (469); Iceland, Malaysia (713); Peru, Jamaica (2023)
$N = 132$. Clustering (day): Portugal, Luxembourg (397); Ireland, Italy, Netherlands, Denmark, Spain, United Kingdom, (457); France, Switzerland (469); Iceland, Malaysia (713); Jamaica, Lebanon, Peru (2023) ; Dominican Republic, Uzbekistan (2071)

Table 4: Parameter estimates: Symmetric spillovers

	(1)	(2)	(3)	(4)
		$N = 103$		$N = 132$
spillover γ	14.883	10.564	11.957	9.225
	(7.465)	(5.386)	(4.482)	(2.913)
constant	10.594	8.910	7.847	6.135
	(0.578)	(0.760)	(0.469)	(0.355)
LNPCY	-0.798	-0.729	-0.535	-0.392
	(0.068)	(0.074)	(0.055)	(0.045)
LNPOP	-0.262	-0.366	-0.235	-0.276
	(0.065)	(0.074)	(0.050)	(0.041)
LNCFC	-0.027	0.114		
	(0.033)	(0.058)		
ART5	-1.211	-1.434	-1.494	-1.034
	(0.372)	(0.408)	(0.300)	(0.173)
LONDON_ART5		1.757	1.801	1.341
		(0.465)	(0.443)	(0.175)
$\lambda * 10^4$	18.353	20.137	19.614	12.844
	(1.164)	(2.352)	(2.129)	(1.023)

Standard errors are in parentheses.

Table 5: Summary statistics: Trade sample

Variable	Obs	Mean	Std. Dev.	Min	Max
DAY	95	1393.516	881.775	197	399
LNPCY	95	7.473	1.663	4.573	10.650
LNPOP	95	2.256	1.574	-1.415	6.972
LNCFC	95	6.015	2.686	0.833	12.631
PRODUCER	95	0.179	0.385	0	1
ART5	95	0.726	0.448	0	1
LONDON_ART5	95	0.484	0.502	0	1
LATDEG	95	23.884	17.627	0	90

Clustering (day): Italy, Netherlands, Denmark, Spain (457); France, Switzerland (469); Iceland, Malaysia (713); Peru, Jamaica (2023)

Table 6: Parameter estimates: Trade spillovers

	(1)	(2)	(3)	(4)	(5)	(6)
	total exports			chemical exports		
spillover γ	0.0001	0.0000	0.0002	0.0002	0.0002	0.0003
	(0.0002)	(0.0089)	(0.0004)	(0.0002)	(0.0001)	(0.0004)
constant	6.443	8.004	3.578	6.723	6.939	5.520
	(1.017)	(0.961)	(1.063)	(0.651)	(0.797)	(0.855)
LNPCY	-0.291	-0.579	-0.049	-0.260	-0.447	-0.096
	(0.122)	(0.114)	(0.081)	(0.083)	(0.101)	(0.080)
LNPOP		-0.292	-0.014		-0.287	0.019
		(0.073)	(0.030)		(0.070)	(0.028)
LNCFC	-0.219		-0.180	-0.252		-0.287
	(0.078)		(0.092)	(0.062)		(0.071)
ART5	-1.488	-1.563	-0.701	-1.513	-1.153	-1.265
	(0.412)	(0.420)	(0.417)	(0.383)	(0.359)	(0.501)
LONDON_ART5	2.402	2.618	2.202	3.107	3.233	3.326
	(0.804)	(0.910)	(0.949)	(0.699)	(0.743)	(0.774)
$\lambda * 10^4$	21.647	22.783	18.943	26.527	26.859	27.670
	(5.721)	(5.998)	(6.459)	(4.605)	(4.608)	(5.037)

$N = 95$. Standard errors are in parentheses.

Table 7: Parameter estimates: Treaty spillovers

	(1)	(2)	(3)
spillovers γ	5.245	3.673	10.311
	(0.217)	(0.745)	(0.113)
constant	3.172	7.890	3.614
	(0.038)	(0.047)	(0.028)
LNPCY	-0.091	-0.729	-0.133
	(0.007)	(0.006)	(0.003)
LNPOP	-0.113	-0.461	-0.158
	(0.006)	(0.007)	(0.003)
LNCFC	-0.039	0.173	
	(0.006)	(0.005)	
ART5	-0.425	-0.900	-0.494
	(0.018)	(0.018)	(0.012)
LONDON_ART5	2.120	2.131	2.173
	(0.018)	(0.0270)	(0.014)
LATDEG		0.012	
		(0.000)	
λ *10^4	18.249	21.733	19.183
	(0.086)	(0.191)	(0.071)

$N = 103$. Asymptotic standard errors are in parentheses.

Table 8: Parameter estimates: CFC spillovers

	(1)	(2)	(3)
spillover γ	14.363	10.383	7.127
	(6.361)	(5.095)	(3.618)
constant	3.338	5.566	6.021
	(0.341)	(0.904)	(0.610)
LNPCY	-0.069	-0.297	-0.402
	(0.028)	(0.089)	(0.070)
LNPOP	-0.123	-0.167	-0.285
	(0.053)	(0.051)	(0.055)
PRODUCER		-0.036	0.376
		(0.023)	(0.203)
ART5	-0.353	-0.912	-0.637
	(0.228)	(0.325)	(0.246)
LONDON_ART5	1.866	1.673	1.983
	(0.704)	(0.583)	(0.577)
LATDEG			0.013
			(0.005)
$\lambda * 10^4$	18.863	18.469	20.710
	(4.105)	(3.930)	(3.486)

$N = 103$. Standard errors are in parentheses.