

International Environmental Agreements with Mixed Strategies and Investment

Fuhai Hong* Larry Karp†

May 21, 2010

Abstract

We modify the canonical participation game used to study International Environmental Agreements, considering both mixed and pure strategies at the participation stage, and including a prior cost-reducing investment stage. The use of mixed strategies at the participation stage reverses a familiar result and also reverses the policy implication of that result: with mixed strategies, equilibrium participation and welfare is higher in equilibria that involve higher investment. Nations' use of mixed rather than pure strategies creates endogenous risk; the equilibrium participation probability increases with risk aversion.

Keywords: International Environmental Agreement, climate agreement, participation game, investment, mixed strategy, risk aversion.

JEL classification numbers C72, H4, Q54

*Department of Economics, the Hong Kong University of Science and Technology, email: fhhong@ust.hk

†Corresponding author: Department of Agricultural and Resource Economics, 207 Giannini Hall, University of California, Berkeley CA 94720, tel 510 642 7199. email: karp@berkeley.edu

1 Introduction

Limiting increases in Greenhouse Gas (GHG) stocks requires an international effort. The global nature of GHG pollution and the fact of national sovereignty complicate the usual problems of eliciting the provision of a public good, in this case, reduction of GHG emissions. Although there are many examples of successful International Environmental Agreements (IEAs), these do not involve problems on the same scale as climate change. In the absence of historical analogies, economic theory plays an especially important role in helping to think about ways of addressing the climate policy problem. Much of this theory relies on a particular type of participation game, in which countries decide whether to join or stay out of an IEA, and then the resulting IEA decides on the level of abatement (Wagner, 2001 and Barrett, 2003). This theory is pessimistic about the equilibrium level of participation. It also implies that actions that might appear to promote membership, such as investment that reduces abatement costs, can be counterproductive.

We use the familiar type of game, but unlike most of the literature we emphasize a mixed strategy rather than a pure strategy equilibrium, and we include a prior stage at which countries individually decide whether to invest in a public good that reduces abatement costs. For most (but not all) levels of abatement cost, a mixed strategy at the participation stage leads to a lower level of expected participation, and lower expected welfare, compared to the pure strategy equilibrium. In this respect, the model with mixed strategies is even more pessimistic than the model with pure strategies. However, the use of mixed strategies creates risk, and the presence of risk aversion increases the probability of participation.

The two participation games, with mixed and with pure strategies, induce two investment games. In general, there are multiple pure strategy equilibria to both of these investment games, but they have markedly different characteristics and different policy implications. With pure strategies at the participation stage, higher investment leads to lower participation and lower welfare for all agents. With mixed strategies at the participation stage, higher investment leads to higher participation and higher welfare for all agents. Due to the existence of multiple Pareto-ranked equilibria, there is a coordination problem in both of the investment games. In the game with pure strategies, society would like to coordinate on the low-investment equilibrium, and in the game with mixed strategies, society would like to coordinate on the high-investment equilibrium.

The principal contributions of this paper stem from our use of mixed strategies. In view of the fact that most of the previous papers that use participation games to model IEAs assume that homogenous agents (countries) use pure strategies, it is important to consider carefully the modeling choice between pure and mixed strategies. The rest of this Introduction discusses this choice. In the remainder of the paper we present the model and then provide the analysis.

With pure strategies, the Nash equilibrium selects the number but not the identity of the participants; ex ante identical agents behave differently. With (symmetric) mixed strategies, ex ante identical agents use the same strategy but they have different outcomes, and the number of participants is the realization of a random variable.

The main argument in favor of modeling the game using pure rather than mixed strategies is that the outcome in the former but not in the latter is “renegotiation proof”. The pure strategy equilibrium size equals the “minimally successful” IEA (one that does “something useful”) and no agent wants to unilaterally change its participation decision. In contrast, in most outcomes to a mixed strategy equilibrium, the number of participants is either smaller or larger than needed to achieve a minimally successful IEA. In all these cases, one or more agent would like to unilaterally change their participation decision.

Renegotiation proofness of the pure strategy Nash equilibrium is not robust and is therefore not so compelling that it should be a requirement of every equilibrium. Renegotiation proofness arises because of the simplicity of the game, and it is easy to overturn. For example, suppose that agents’ abatement costs are imperfectly correlated random variables that are realized after the participation decisions, and at the time of the participation decision all agents face the same distribution of costs. In addition, suppose that costs enter the payoffs linearly and costs are not verifiable, so the IEA cannot condition its instructions on the realizations. Under these assumptions, the pure strategy equilibria to the participation game with random costs and to the game with fixed costs (where the two costs are equal in expectation) are the same, but the pure strategy equilibrium in the game with random costs may not be renegotiation proof.

Alternatively, agents may have more sophisticated beliefs about how their provisional decision to join or leave a coalition would affect other agents’ participation decisions (Chwe, 1994, Xue, 1998 and Ray and Vohra, 2001). Diamantoudi and Sartzetakis (2002), Eyckmans (2001), de Zeeuw (2008)

and Osmani and Tol (2009) use models of this sort to study IEAs. With these sophisticated beliefs, the pure strategy Nash equilibrium may not be subgame perfect, and therefore not renegotiation proof.

Few papers in the IEA literature examine mixed strategies. Kohnz (2006) considers a mixed strategy in a game with a nonstandard structure in which the IEA may behave sub-optimally conditional on the number of members. An earlier stage determines the minimum number of members needed for ratification. Sandler and Sargent (1995) and McGinty (2010) use mixed strategies in one-stage coordination games. The “integer problem”, which is important in standard IEA models and in our’s, is ignored in McGinty (2010) and irrelevant in Sandler and Sargent (1995).

The general literature on participation games routinely uses mixed strategies. Our paper is most closely related to Dixit and Olson (2000), which has a similar structure to Palfrey and Rosenthal (1984). Dixit and Olson rely primarily on numerical examples to characterize the mixed strategy equilibrium, whereas we obtain analytic results. More importantly, the two papers have different objectives. Dixit and Olson’s primary point is that Coasian bargaining does not solve the participation problem; they demonstrate this by showing that in general participation is low even though the resulting members of the agreement engage in Coasian bargaining. We find that in special circumstances Coasian bargaining *is* consistent with high participation, and that investment in low cost abatement technology may lead to exactly these circumstances. As we noted above, a model that uses pure strategies typically concludes that such investment is inimical to high levels of participation.

A coordination problem arises with pure but not with mixed strategies: How does the group choose which agent participates and which stays outside the agreement under pure strategies? The coordination difficulty also motivates the use of mixed strategies in the corporate takeovers literature (Bagnoli and Lipman, 1988; Holmstrom and Nalebuff, 1992) and wars of attrition (Maskin, 2003).

Crawford and Haller (1990) and Makris (2009) provide additional support for the choice of the symmetric mixed-strategy equilibrium. Crawford and Haller (1990) consider games in which identical players have identical preferences among multiple equilibria, but must learn to coordinate over which equilibrium to play by repeatedly playing the game. They show that in finite time the equilibrium converges to a strategy profile in which all players obtain equal payoffs. In the IEA game only the symmetric mixed-strategy

equilibrium has this property. Harsanyi (1973) and Makris (2009) point out that the symmetric mixed-strategy equilibrium of the game with no inherent uncertainty can be viewed as an approximation to a model where there is small uncertainty about the players' preferences and the expected group-size is sufficiently large.

An equilibrium concept can also be evaluated based on the economic intuition it produces. The pure strategy equilibrium provides interesting insight, by explaining why a low abatement cost (or more generally, a high potential gain from cooperation) leads to low participation: the amount of participation needed for a minimally successful IEA is lower with low abatement costs, and the equilibrium membership in pure strategies exactly equals this level of participation. The mixed strategy equilibrium reinforces this pessimistic insight, because in most circumstances it results in even lower expected membership than in the pure strategy equilibrium.

The new insight is that for sufficiently low abatement costs, the loss to an individual country of failing to achieve a minimally successful IEA is large, but the cost to that country of being "superfluous" (in a sense made precise below) is small. In this circumstance, a country is indifferent between joining and staying out of the IEA, as the mixed strategy equilibrium requires, only if the probability that it is superfluous is large. That large probability means that the equilibrium probability that other countries join is large, which in turn means that the expected level of membership is large. In this respect, the mixed strategy equilibrium is richer than the pure strategy equilibrium, because the former yields a more diverse set of outcomes, and provides clear intuition about the fundamentals that determine the nature of the outcome. More significantly, as noted above, the two types of strategies lead to the opposite policy recommendation regarding the desirability of investment that reduces abatement costs.

In addition, risk aversion has different implications when agents use pure or mixed strategies. Boucher and Bramouille (2009) show that in a pure strategy equilibrium, risk aversion together with *exogenous* uncertainty about environmental damages may weakly decrease equilibrium participation. Our Appendix explains that the effect (on participation in a pure strategy) of risk aversion with exogenous risk is ambiguous. Absent exogenous risk, countries face no risk in a pure strategy Nash equilibrium to the participation game, so risk aversion has no effect on the equilibrium. With mixed strategies, in contrast, there is endogenous risk at the participation stage. An increase in risk aversion always increases participation probabilities, and the effect can

be large for high levels of risk aversion.

In fact, both types of risk are important for environmental problems such as climate change. There is tremendous uncertainty about the costs and benefits of reducing emissions, and there is also uncertainty about the ability of nations to negotiate a climate agreement that leads to meaningful action. These two types of uncertainty may have opposite effects on equilibrium expected participation. A pure strategy equilibrium does not capture the uncertainty about the ability of nations to agree on a meaningful treaty; a mixed strategy equilibrium captures exactly this uncertainty. Benedick (2009, page xv), in discussing his experience as a US negotiator in many IEAs, objects to what he considers academics' tendency to view the negotiating process as mechanistic, yielding a deterministic outcome. He emphasizes the contingency of the process, and the possibility of surprises. The empirical literature that attempts to explain participation in IEAs (Murdoch, Sandler and Vijverberg, 2003) of course uses a probabilistic model.

Some IEAs, including the Montreal Protocol, appear to be more consistent with a mixed than a pure strategy equilibrium. Participation costs in the Montreal Protocol were small and the benefit of cooperation large (Barrett 2003, Chapter 8). In this circumstance, the pure strategy equilibrium predicts low membership, but in some cases the mixed strategy equilibrium predicts high *expected* participation, and in all cases it is consistent with high *actual* participation. The Montreal Protocol has high participation.

More generally, some scholars claim that international environmental agreements do not result in significant change relative to non-cooperative behavior, so these IEAs achieve little; see Downs et al. (1995) and Barrett (2003) for an assessment of this view. Decision-makers' actions depend on their perception of the private cost-benefit ratio of a project that provides a public good. Economists' point estimate of that perceived cost-benefit ratio may be only slightly above 1, and we may also observe high participation in an IEA. In this case, perhaps decision-makers' perceived cost-benefit ratio is actually less than 1, and the IEA really does not alter non-cooperative behavior and so achieves nothing. Another explanation is that the perceived cost-benefit ratio is only slightly greater than 1 and decision-makers use mixed strategies. In this case the IEA achieves a great deal, because equilibrium participation in the IEA is high but in the absence of the IEA decision-makers would choose not to provide the public good.

2 The model

Our model differs from the canonical IEA participation game in two respects. We examine mixed rather than pure strategies at the participation stage, and we include a prior stage at which countries decide whether to make an investment that reduces all countries' abatement costs. There are two public goods, investment and abatement. We have in mind investment in developing a technology that can be easily replicated. Patent law enables investors to capture some of the benefit of their investment, but in many cases there are significant spillovers. Our emphasis on the situation where investment is a pure public good captures an important element of the problem.

Modeling investment as a public good means that agents that did invest in the first stage and those that did not invest have the same abatement costs at the participation stage. Consequently, we can consider symmetric strategies at the participation stage, a fact that greatly simplifies the analysis. The public good nature of investment also means that for the world as a whole, investment has a high return.

Golembek and Hoel (2005, 2006, 2008) consider the effect of investment with technological spillovers when membership in the IEA is exogenous; in our setting, investment affects IEA membership. Barrett (2006) studies the situation where countries choose investment non-cooperatively prior to the participation game, at which stage countries use pure strategies. He points out that investment opportunities do not necessarily help to solve the participation problem; we show that investment opportunities can lower welfare in the model with pure strategies. Hoel and de Zeeuw (2009) assume that the IEA chooses both investment and the abatement level after the participation stage, and they allow for the possibility that investment can drive the abatement cost to a level below the threshold at which abatement is a dominant strategy; we exclude that possibility and we adopt the same timing as in Barrett (2006).¹

¹If investment reduces abatement costs only in the country where the new technology is installed, rich countries might want to install the new technology in their own countries and also subsidize the installation of this technology in developing countries. Non-cooperative behavior at the investment stage would be unlikely to lead to the provision of this subsidy, which would instead require a collective agreement at the investment stage, . This alternative model with private investment has the flavor of international climate negotiations, where developing countries seek subsidies from rich countries in order to reduce the cost of transition to a low-carbon economy.

In our model there are N identical countries, each of which has three sequential binary decisions: investment, participation in an IEA, and abatement. Abatement costs and the benefit of abatement are both linear, so in equilibrium each country abates at the level 0 or at capacity, normalized to 1. Given linearity, there is no additional loss in generality in assuming that the abatement decision is binary. The IEA instructs all members whether to abate, and non-members follow their individually rational actions.

We normalize the benefit of each unit of abatement, to each country, to 1. Each country's abatement cost is c , with $1 < c < N$. The first inequality means that it is a dominant strategy for a country acting alone not to abate and the second inequality means that the world is better off when a country abates. In view of our normalization of benefit, c also equals the cost-benefit ratio of abatement.

Conditional on membership m , the IEA instructs its members to abate if and only if $m - c \geq 0$. The function $f(x)$ returns the smallest integer not less than x . There are many pure strategy Nash equilibria, but the unique pure strategy equilibrium with positive abatement, and the only one that we consider hereafter, is $f(c)$. For example, with $c = 3.2$, there are four members in this pure strategy equilibrium. Equilibrium global welfare here is $(N - c) f(c)$. A small reduction in c can lead to a discrete reduction in equilibrium membership, reducing global welfare. The “minimally successful” IEA contains $f(c)$ members; larger IEAs achieve more abatement and higher global welfare, but smaller IEAs achieve nothing. The pure strategy equilibrium equals the minimally successful IEA; at any outcome with more than $f(c)$ members, the “extra” members would want to leave, so an outcome with more than $f(c)$ members is not an equilibrium in pure strategies.

At the participation stage c is given, but previous investment determines its level. If k countries invested, $c = c(k)$, a decreasing function. We will be especially interested in situations where $c(N)$ is close to (but slightly larger than) 1. In this case, if N countries invest, equilibrium membership in the pure strategy equilibrium is 2, and few of the potential benefits of cooperation are realized. The cost of a unit of investment is ε , which might be either large or small.

We first study the participation stage, taking k (and thus $c = c(k)$) as given. Here our goal is to characterize the equilibrium mixed strategy, $p = p(c, N)$, the probability that any country joins the IEA. At this stage all countries understand that the IEA will instruct its members to abate if and only if the number of members is at least $f(c)$. We then study the

investment stage, where each country decides whether to invest and incur the cost ε . Countries have rational expectations, so they understand that their investment decision affects c and thereby affects equilibrium expected participation. The assumption that investment is a public good means that both non-investors and investors have the same probability of participation.

3 The participation stage

We emphasize the risk neutral case. Here, for most but not all of parameter space, the model with mixed strategies is more pessimistic than the model with pure strategies. This result is interesting as a robustness check to one of the major insights of this literature: optimal behavior on the part of the IEA does little to resolve the problem of providing a global public good. We then briefly consider the role of risk aversion. We find that risk aversion always increases equilibrium participation probabilities. For sufficiently high risk aversion, countries are likely to join the IEA.

The IEA instructs its members to abate if and only if there are at least $f = f(c)$ members. Therefore, a country is pivotal if and only if exactly $f(c) - 1$ other countries join. If fewer than $f(c) - 1$ other countries join, then there would still be too few members to elicit abatement even if an additional country joins. If more than $f(c) - 1$ other countries join, then membership by an additional country is (from its own standpoint) superfluous: by joining it has no effect on other countries' abatement but it incurs the net cost $c - 1 > 0$.

With the probability of each country joining equal to p , the probability that a country is pivotal is

$$g(p) = \frac{(N-1)!}{(f-1)!(N-f)!} p^{f-1} (1-p)^{N-f}.$$

The probability that at least f other countries joined is

$$G(p) = \sum_{i=f}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i}.$$

3.1 Risk neutrality

Under risk neutrality, the net expected benefit of joining is the benefit of joining when the country is pivotal, $f - c$, times the probability that it is

pivotal, g , minus the loss when the IEA would have abated even had this country not joined, $c - 1$, times the probability of that event, G . In a mixed strategy equilibrium, p must be such that a country is indifferent between joining and not joining. The equilibrium condition for p is therefore

$$(f - c)g(p) = (c - 1)G(p). \quad (1)$$

Remark 1 *When c is an integer greater than 1, $f(c) = c$ and the left side of this equation is 0. The right side equals 0 if and only if $p = 0$. Thus, $p = 0$ is the only mixed strategy equilibrium when c is an integer.*

For non-integer values of $c > 1$, there are two solutions to equation (1), $p = 0$ and $0 < p < 1$. We are interested in the latter. (The literature using pure strategy equilibria also ignores the trivial pure strategy equilibrium in which no country joins.) To study non-integer values of c we rearrange equation (1) to obtain

$$\frac{f - c}{c - 1} = \frac{G(p)}{g(p)}. \quad (2)$$

When $p \neq 1$ and $p \neq 0$, the right hand side of equation (2) is well defined because $g(p) \neq 0$. We simplify the right side of equation (2) using

$$\frac{G(p)}{g(p)} + 1 = (f - 1) \int_0^1 (1 - t)^{f-2} \left(1 + t \frac{p}{1-p}\right)^{N-f} dt \equiv \sigma(p; N, f). \quad (3)$$

The appendix contains the derivation of this equation and proofs of all propositions.

Equation (3) leads to a simple demonstration of the following:²

Proposition 1 *(i) When c is a non-integer, there is a unique nontrivial mixed strategy equilibrium $p(c, N) > 0$. (ii) Over any interval of c for which f is constant, the equilibrium mixed strategy p is a decreasing function of c .*

²In Palfrey and Rosenthal's (1984) "refund case", the probability of joining is monotonically decreasing in the cost of contribution (Corollary 5.1). In their model, the benefit of being pivotal is fixed at 1 while in our model the benefit of being pivotal is $f - c$. When $f = c$, the benefit is 0 and thus $p = 0$. However, as c increases slightly, $f - c$ jumps to almost 1. This difference explains why the relation between p and c is monotone in their model whereas it is a saw-tooth in our model.

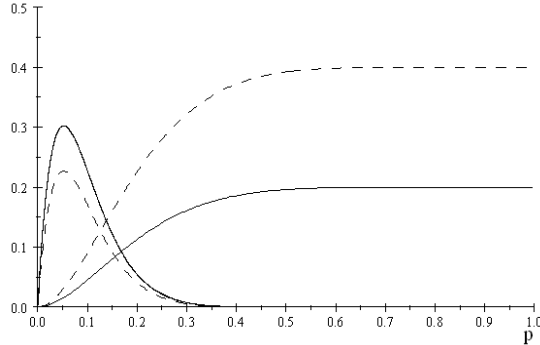


Figure 1: Monotonic curves show $(c - 1)G$ and hump-shaped curves show $(f - c)g$. Solid curves correspond to $c = 1.2$ and dashed curves correspond to $c = 1.4$. $N = 20$ for both pairs of curves.

Figure 1 illustrates the comparative statics of p with respect to c , holding $f(c)$ constant. The monotonic curves graph the right side of equation (1) and the hump shaped curves graph the left side. The solid curves correspond to $c = 1.2$ and the dashed curves correspond to $c = 1.4$, both with $N = 20$. The figure shows that an increase in c , holding $f(c) = 2$, decreases the left side and increases the right side of the equilibrium condition, thereby reducing the equilibrium value of p .

Remark 1 and Proposition 1 imply that the graph of p as a function of c is a saw-tooth, with points of discontinuity from the right occurring at integers. Figure 2 illustrates this relation for $N = 20$ and $c \in (1.1, 6)$. In the pure strategy equilibrium m denotes the deterministic number of participants; here we use m to denote the *expected* number of participants: $m \equiv pN$. Because m is proportional to p , expected membership is also a saw-toothed function of c .

The next result compares the equilibrium expected membership under mixed strategies with membership in the (non-trivial) pure strategy equilibrium:

Proposition 2 (i) For any $c > 2$, the expected membership in the mixed-strategy equilibrium, m , is less than 3. (ii) In the interval $c \in (1, 2]$, m is greater than 2 (the membership of the corresponding nontrivial pure-strategy equilibrium) for c sufficiently small. Moreover, as c approaches its lower bound 1, p converges to 1, and m converges to N .

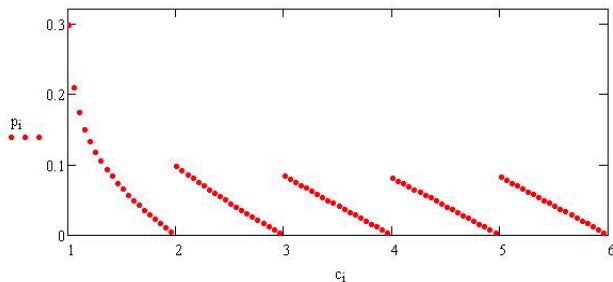


Figure 2: The relation between p and c with $N = 20$ and $c \in (1.1, 6)$.

Barrett (2005) shows that in games with specific functional forms for costs and damages, and continuous action space, the equilibrium number of members is never greater than 3. Proposition 2.i shows that an analogous result holds in the game with binary actions and mixed strategies. As we noted above, a standard result in IEA theory based on pure strategy equilibria in the binary action game is that the fraction of countries that participate in equilibrium, $\frac{f(c)}{N}$, is high if and only if the potential global benefit of cooperation, $N(N - c)$, is small; this result arises because equilibrium membership equals the size of the minimally successful IEA, which is weakly increasing in c and is always less than $c + 1$. The saw-toothed relation between p and c under mixed strategies implies that higher costs have an ambiguous effect on expected participation when countries use mixed strategies. When c increases, holding $f(c)$ constant, the probability of participation and expected participation decrease; however, if c increases from slightly below to slightly above an integer, the probability of participation and expected participation both increase. Moreover, we find that for $c \geq 2$ expected participation is lower under mixed strategies compared to pure strategies. In this respect, the model with mixed strategies is even more pessimistic than the model with pure strategies.³

³We also calculated expected membership, conditional on membership being greater than or equal to $f(c)$, i.e. conditional on the IEA abating. For $N = 20$ we found that for $c > 2$, this conditional expectation is always between $f(c)$ and $f(c) + 1$. Thus, for $c > 2$, the use of mixed rather than pure strategies makes an IEA with at least $f(c)$ members unlikely, and even when such an IEA arises, the expected level of membership is close to $f(c)$.

An important difference is that when abatement costs are sufficiently low, expected participation is high when countries use mixed strategies. The explanation is straightforward. If c is close to 1, a country's loss when it joins the coalition "unnecessarily" ($c - 1$) is small. However, it loses nearly one unit if the membership does not reach the critical size of 2. Therefore, for small values of c a country is indifferent between joining and staying out, only if it is very confident that the IEA will reach its critical size, i.e. when p is close to 1.

The equilibrium p is close to 1 only as the cost-benefit parameter, c , gets close to the threshold 1, below which abatement is a dominant strategy. For example, for $N = 20$, in equilibrium $p > 0.5$ if and only if $c - 1 < 0.18$, and $p > 0.9$ if and only if $c - 1 < 8.1 \times 10^{-8}$. If one interprets this model literally, the conclusion is that for almost all parameter space the outcome with pure strategies provides an optimistic upper bound to the outcome when agents use mixed strategies. Thus, the results here provide a robustness check for the pessimistic result that optimal behavior by the IEA does little to solve the problem of providing a global public good. Of course, it may be a mistake to take a literal interpretation of a model that is intended to provide qualitative insights. The new qualitative insight is that there is a special case in which optimal behavior by the IEA *does* essentially solve the problem of providing the public good. However, the real interest in mixed strategies is that they produce qualitatively different insight about the role of investment, as we discuss in Section 4.

In the wake of the negotiations at Copenhagen 2009, some people believe that a smaller group of countries will have a better chance of achieving a successful climate agreement. Of course, the model with homogenous agents ignores the heterogeneity that actually exists among countries, one of the most salient features of the climate problem. Nevertheless, it is interesting to note that an increase in the number of agents does reduce the equilibrium probability that any country will participate, and has ambiguous effects on the expected level of participation. The following proposition summarizes these comparative statics.

Proposition 3 (i) *The equilibrium probability of joining, p , decreases with the number of countries, N .* (ii) *For N sufficiently small, $\frac{dp}{dN} < 0$, and for N sufficiently large $\frac{dp}{dN} > 0$.* (iii) *When c converges to 1 from the right, $\frac{dp}{dN} > 0$.* (iv) *When c converges to f from the left, $\frac{dp}{dN} < 0$.*

The intuition for the first part of the proposition is that, for a given p , the probability that a country is pivotal decreases in N , and the probability that more than $f - 1$ other countries join increases with N . Therefore, for a given p , the expected net benefit of joining decreases with N . In order to maintain indifference, the value of p must decrease. It is not surprising that expected membership, pN , is non-monotonic in N : a larger N increases one factor in this expectation and decreases the other factor.

The next proposition gives the formula for the equilibrium expected participation as $N \rightarrow \infty$.⁴

Proposition 4 *As $N \rightarrow \infty$, (i) $p \rightarrow 0$ and (ii) $m \rightarrow \lambda$, where λ is determined by*

$$\frac{1}{c-1} = \int_0^1 (1-t)^{k-1} e^{t\lambda} dt; \quad (4)$$

and (iii) membership follows a Poisson distribution with parameter λ .

The right side of equation (4) increases in λ while the left side decreases in c . Therefore, over a region where k remains constant, the equilibrium λ decreases in c . Where c passes through an integer value, the right side jumps down for given λ so the equilibrium λ must jump up. Thus, in the limit as $N \rightarrow \infty$ we have the same kind of saw tooth pattern as in the case with finite N .

3.2 Risk aversion

Beginning with the deterministic model and pure strategies, there are two ways to introduce risk. In moving from pure to mixed strategies with no exogenous uncertainty, we create *endogenous* risk, specifically, risk about participation. In contrast, if we maintain the assumption of pure strategies but make the cost-benefit ratio c stochastic, we introduce *exogenous* risk. Depending on the manner in which exogenous risk is introduced, an increase in risk aversion can either increase or decrease equilibrium participation in a pure strategy equilibrium. Exogenous risk is incidental to our objectives, so we relegate a discussion of it to the Appendix. This section shows that with the endogenous risk arising from mixed strategies, an increase in risk aversion always increases equilibrium expected participation.

⁴Myerson (1998a, 1998b, 2005) uses the Poisson limit theorem to study the limiting behavior in binary action games when the number of players approaches infinity.

Define A as a country's baseline income in the absence of environmental damage or abatement, and let $y^s(m)$ and $y^n(m)$ be, respectively, income of a signatory and a non-signatory to an IEA with m members.

$$y^s(m) = \left\{ \begin{array}{l} A - (N - m + c) \text{ if } m \geq c \\ A - N \text{ if } m < c \end{array} \right\}$$

$$y^n(m) = \left\{ \begin{array}{l} A - (N - m) \text{ if } m \geq c \\ A - N \text{ if } m < c \end{array} \right\}.$$

Let $U(y)$ be a concave transformation, strictly concave in a subset of the interval $(A - N + f - c, A)$. With preferences $U(y)$, countries are risk averse, and strictly risk averse over some interval. Replacing preferences by $V(U(y))$, where V is concave, represents an increase in risk aversion. We have the following:

Proposition 5 (i) *The introduction of risk aversion (changing the payoff from y to $U(y)$) increases the equilibrium participation probability.* (ii) *An increase in risk aversion increases the equilibrium participation probability.*

The intuition for the proposition is straightforward. Note that for both signatories and non-signatories, income is (weakly) increasing in the number of members. Under risk aversion, the marginal utility of income decreases with the level of income. A country loses $c - 1$ units of income by joining if at least $m \geq f$ other countries have joined. The larger is m , the larger is income, and therefore the smaller is the utility loss of joining “unnecessarily” (i.e., when $m \geq f$ other countries have joined). A country gains $f - c$ by joining only if $m = f - 1$, i.e. when income is low, and the marginal utility of income is relatively high. Risk aversion therefore increases the utility gain of joining when $m = f - 1$ and decreases the utility loss of joining when $m \geq f$. In order for a country to remain indifferent between joining and not joining, the probability of $m \geq f$ must increase. That cumulative probability increases if and only if p , the probability that an individual country joins, increases. This intuition holds even if actions are continuous (as they would be under exogenous risk about c) rather than binary.

Thus, in the presence of endogenous risk, an increase in risk aversion increases the equilibrium expected size of the IEA; the effect of exogenous risk (in a pure strategy equilibrium) depends on specifics of the model, as we show in the Appendix. The climate problem of course has both types of risk: we

do not know whether nations will agree to a treaty that results in meaningful actions, and if such a treaty does emerge, its net benefits are stochastic. We can easily include exogenous risk (about c or about the damage parameter) in our model. However, the combination of both exogenous and endogenous risk appears to add little insight, above that obtained by considering both types of risk in isolation. To assess the magnitude of the effect of risk aversion we consider an analytic example here. The appendix shows that a model with constant relative risk aversion yields similar results.

Example: piece-wise linear utility Suppose that utility is piece-wise linear in income, with the kink at $A - N + f - c$, in order to satisfy the assumption that the country is strictly risk averse over a subset of the interval $(A - N + f - c, A)$:

$$U = \left\{ \begin{array}{l} y \text{ if } y \geq A - N + f - c \\ \frac{-2w}{1-2w} (A - N - c + f) + \frac{1}{1-2w} y \text{ if } y < A - N + f - c \end{array} \right\}, \quad (5)$$

for $w \in [0, 0.5)$. With this utility function, the marginal utility of income equals 1 for $y \geq A - N + f - c$ and marginal utility equals $\frac{1}{1-2w}$ for $y < A - N + f - c$. As w increases from 0 to its supremum value of 0.5, marginal utility for $y < A - N + f - c$ increases from 1 to infinity.

The parameter w is a measure of risk aversion in the neighborhood of the kink in the utility function.⁵ The fact that utility is linear above and below the kink, $A - N + f - c$, means that we are able to use the argument that led to the equilibrium condition (2) to show that the equilibrium condition is now

$$\frac{\frac{1}{1-2w} (f - c)}{c - 1} = \frac{G(p)}{g(p)}. \quad (6)$$

The numerator of the left side is the marginal utility of the loss when the country decides not to join and there are $f - 1$ members, and the denominator

⁵To see this, suppose that income were a random variable that takes the value $A - N + f - c - \varepsilon$ or $A - N + f - c + \varepsilon$, each with probability 0.5. Let the risk premium r equal the amount that society would pay to stabilize income at its expected value, $A - N + f - c$, leaving society with the sure income $A - N + f - c - r$. A calculation shows that $w = \frac{r}{\varepsilon}$, so w is the risk premium as a fraction of the random shock ε . Under risk neutrality ($w = 0$) the risk premium is 0, and as society becomes infinitely averse to risk ($w \rightarrow 0.5$) the risk premium approaches 0.5ε . In a general model with strictly concave utility, the supremum of $\frac{r}{\varepsilon}$ is 1.

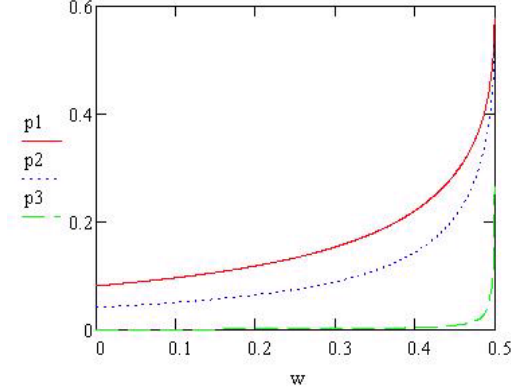


Figure 3: Relation between w and equilibrium p for $N = 20$ and $c = 5.01$ (solid), $c = 5.5$ (dotted) and $c = 5.99$ (dashed)

is the marginal utility of the loss when country decides to join and there are $m \geq f$ other members. An increase in the risk aversion parameter w increases the left hand side, due to the increase in $\frac{1}{1-2w}(f-c)$. The fact that the right side of equation (6) is increasing in p implies that the equilibrium p increases in w . The fact that $\frac{G(p)}{g(p)} \rightarrow \infty$ as $p \rightarrow 1$ also implies that the equilibrium participation probability approaches 1 as the risk aversion parameter approaches its supremum, $w = 0.5$.

Figure 3 shows the relation between the equilibrium p and the exogenous parameter $w \in [0, 0.4995)$ for $N = 20$ and for three values of $c \in \{5.01, 5.5, 5.99\}$. Under risk neutrality ($w = 0$) the participation probability is less than 0.1 for all three values of c . The participation probability approaches 1 as $w \rightarrow 0.5$.

4 The investment stage

Here we investigate how, under risk neutrality, the use of pure or mixed strategies in the participation game affects the equilibrium in the investment game, and overall equilibrium welfare. By incurring the investment cost ε , a country reduces the abatement cost in the next stage. We emphasize the situation where countries use pure strategies at the investment stage.

The two participation games (with pure strategies and with mixed strate-

gies) induce two investment games. In both games, there may be multiple equilibria, but the nature of the multiplicity is quite different. When countries use pure strategies at the participation game, the multiplicity at the investment stage arises from the non-monotonicity of payoffs in costs. In this game, simulations show that the equilibrium with higher investment leads to lower welfare, exclusive of investment costs. Because the countries that invest incur the investment cost, all countries have lower welfare in the high-investment equilibrium, so the low-investment equilibrium is Pareto dominant.

The investment game when countries use mixed strategies at the participation stage is more interesting. The fact that p , and therefore $m = pN$, is a saw-toothed function of c means that the reduction in c caused by higher investment might lower equilibrium participation and hence lower welfare, just as with pure strategies. However, Proposition 2.ii implies that if the investment lowers abatement costs sufficiently, the benefit of investment may be high. In this case, the investment decisions may be strategic complements. Consider matters from the standpoint of a country that has not yet decided whether to invest. If that country believes that few other countries will invest, it may rationally calculate that its own investment would lower expected participation in the next stage and could be unprofitable. However, if it believes that many other countries will invest, the additional decrease in abatement costs caused by its investment might then increase expected membership and be profitable. In this case, the mixed strategy at the participation game induces a coordination game at the investment stage, with high and low investment equilibria. Here, the high investment equilibrium is likely to be Pareto-dominant.

The non-monotonicity of payoffs with both pure and with mixed strategies, and the fact that they are defined only over integer values of k , complicates the analysis. However, the basic conclusions are intuitive, and we illustrate these using a numerical example. Additional simulations confirm that the reported results are robust.

4.1 Pure strategies at the participation stage

Investment is a pure public good, so a country's investment decision does not distinguish that country at the participation stage. As in Barrett (2006) and Rubio and Ulph (2007), we therefore assume that all countries expect to face the same probability that they will be one of the participants in the

next stage, regardless of whether they invested. Therefore, the probability that any country joins the IEA in the pure strategy equilibrium is simply the membership of the nontrivial pure strategy participation, $f(c)$, divided by the total number of countries, N . Let $\varpi(k)$ be the next-stage expected payoff given investment level k :

$$\varpi(k) = f(c(k)) - \frac{f(c(k))}{N}c(k).$$

The benefit of investment for a country given that $k - 1$ other countries invest is the change in its next-stage expected payoff:

$$\phi(k - 1) = \varpi(k) - \varpi(k - 1).$$

Given the belief that $k - 1$ other countries invest, a country will invest if and only if $\phi(k - 1) \geq \varepsilon$.

If $\phi(0) < \varepsilon$, then there is a corner equilibrium in which no country invests; if $\phi(N - 1) \geq \varepsilon$, there is a corner equilibrium in which every country invests. If there exists a K , with $0 < K < N$ and $\phi(K - 1) > \varepsilon > \phi(K)$, then there is an interior equilibrium in which K countries invest and $N - K$ countries do not invest.

Global welfare as a function of k is

$$\Phi(k) = (N - c(k))f(c(k)) - k\varepsilon.$$

A larger value of k lowers abatement costs but it increases total investment costs and might decrease participation; therefore, global welfare may be non-monotonic in k . The difference in global welfare under zero investment and under a positive value of k is

$$\begin{aligned} \Delta(k) &\equiv (N - c(0)) [f(c(0)) - f(c(k))] - f(c(k)) [c(0) - c(k)] + k\varepsilon \geq \\ &N(c(0) - c(k) - 1) - (c(0) - c(k))(c(0) + c(k)) + c(k) + k\varepsilon. \end{aligned}$$

A sufficient condition for investment to change the level of participation is $c(0) - c(k) \geq 1$. When this condition holds, $\Delta(k) > 0$ for sufficiently large N . In this situation, global welfare is lower under $k > 0$ compared to $k = 0$.

The following example illustrates some of the possibilities:

$$c(k) = \frac{5.99}{0.248k + 1}, \quad N = 20, \quad \varepsilon = 0.11 \quad (7)$$

With this example, unit abatement costs fall from 5.99 to 1.005 and participation in the pure strategy equilibrium falls from 6 members to 2 members as the amount of investment, k ranges from 0 to 20.

Table 1 shows values of the net benefit of investment, $\phi(k-1) - \varepsilon$, and global welfare, Φ , for several values of k ; the asterisks identify the equilibrium values of k . For this example $k = 0$ and $k = 2$ are the only two equilibria. If no other countries invest, it does not pay any country to invest ($\phi(0) - \varepsilon < 0$). If one other country invests, it pays any country to invest ($\phi(1) - \varepsilon = 0.09 > 0$), but once two (or more) countries invest, further investment yields a negative payoff. Thus, $k = 2$ is the only interior equilibrium. In this example, both investors and non-investors have a higher payoff when 0 rather than 2 countries invest, so the non-investment equilibrium is Pareto-dominant.

k	$\phi(k-1) - \varepsilon$	$c(k)$	m	$\Phi(k)$	$\varpi(k)$
0*	NA	5.99	6	84.1	4.20
1	-0.51	4.80	5	75.9	3.8
2*	0.09	4.004	5	79.8	4.00
3	-0.80	3.43	4	65.9	3.31
4	-0.02	3.01	4	67.5	3.40

Table 1: Investment, membership and welfare with pure strategies at the participation stage. Equilibrium k denoted by *.

This example also shows that the multiplicity of equilibrium is due to the non-monotonicity of the payoffs in c , and disappears for slightly larger investments costs. Here, for $\varepsilon > 0.11 + 0.09 = 0.2$ there is a unique equilibrium, $k = 0$. The example also shows that if investment were determined by a social planner who maximizes world welfare (rather than as a non-cooperative equilibrium), the planner would choose 0 investment.

Figure 4 shows the benefit to investment as a function of $k-1$. The function is non-monotonic; however, because an interior equilibrium k requires $\phi(k-1) > \varepsilon > \phi(k)$, only values of k for which $\phi(k-1) > 0$ are candidates for an equilibrium. For the set where $\phi(k-1) > 0$, the function $\phi(k-1)$ is decreasing in k . This fact means that there is at most one interior equilibrium, and an increase in as we increase ε from 0, the interior equilibrium has a smaller k .

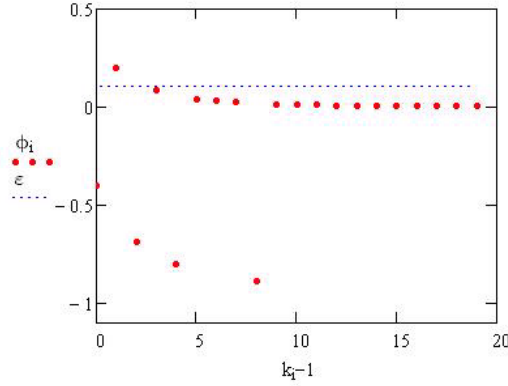


Figure 4: The benefit of investment under pure strategies

4.2 Mixed strategies at the participation stage

The equilibrium p is a function of c , which depends on the level of investment, k , so the equilibrium p is a function of k . Conditional on k , a country's expected welfare equals the probability that it participates times its expected payoff conditional on participation, plus the probability that it does not participate times its expected payoff conditional on non-participation. In equilibrium, these two conditional expectations are equal, so we can write the unconditional expectation as

$$w(k) = \sum_{i=f}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i} i, \quad (8)$$

which is the expected abatement when at least one country does not join. When $k-1$ other countries invest, the benefit to a country of also investing is

$$\Omega(k-1) = w(k) - w(k-1).$$

If $\Omega(0) < \varepsilon$, then there is a corner equilibrium in which no country invests; if $\Omega(N-1) \geq \varepsilon$, there is a corner equilibrium in which every country invests; again, there may be interior equilibria.

Proposition 2 states that when $c \in (1, 2]$, p decreases in c and converges to 1 as c approaches 1. Define $\hat{c} \in (1, 2)$ as the value of c that induces expected participation under mixed strategies equal to the pure strategy level of participation when there is no investment; \hat{c} is the solution to $p(\hat{c}) =$

$\frac{f(c(0))}{N}$. The following proposition considers the case where a high level of investment leads to large cost reductions, which may then lead to relatively high participation and abatement. In this situation, for sufficiently high investment (large k) each country's expected payoff is an increasing function of aggregate investment.

Proposition 6 *Assume that N is sufficiently large and c decreases sufficiently rapidly in k that there exists a $\hat{k} < N$ at which $c(\hat{k}) \leq \hat{c}$. In this case, for any $k_1 \geq \hat{k}$ and all $k_2 < k_1$, $w(k_1) > w(k_2)$.*

This proposition implies that welfare at the participation stage (i.e., ignoring investment cost) is maximized when all countries invest.

Interior equilibria might be eliminated by increasing ε , for the same reason as with pure strategies. For example, if $\varepsilon \geq w(k) - w(k-1)$ for any k such that $c(k) > \hat{c}$, then $\varepsilon > w(k) - w(k-1)$; in this case, there can be no interior equilibria with investment less than \hat{k} . By eliminating equilibria, a large value of ε eases the coordination problem. Of course, an extremely large value of ε eliminates all equilibria with positive investment. However, if

$$\Omega(N-1) > \varepsilon > \Omega(k-1) \text{ for all } k < \hat{k}$$

then there are two equilibria, at which $k = 0$ and $k = N$. (There may be other equilibria with investment $\hat{k} < k < N$.) If it is individually rational for a country to invest, then that investment also increases global welfare. We saw that with pure strategies at the participation stage, it might be individually rational to invest in cases where this lowers social welfare. The expected global welfare as a function of k is

$$\Psi(k) = Nw(k) - k\varepsilon.$$

Figure 5 shows that (under second stage mixed strategies) the benefit to investment, although non-monotonic, increases rapidly for large k . The figure also shows lines at three values of ε . An interior equilibrium requires $\Omega(k) > \varepsilon > \Omega(k-1)$. Increases in ε eliminate the smaller equilibria, but not the boundary equilibrium $k = N$ until ε becomes large.

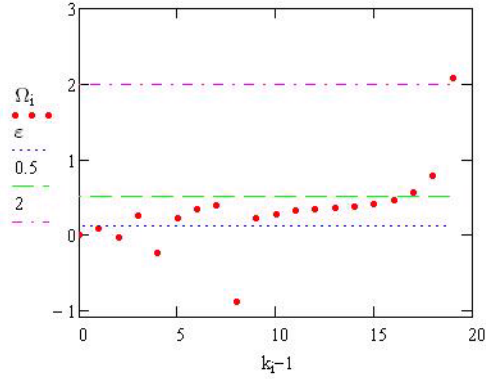


Figure 5: The benefit to investment $\Omega(k)$ and three levels of costs, $\varepsilon \in \{0.11, 0.5, 2\}$

k	$\Omega(k-1) - \varepsilon$	$c(k)$	m	$\Psi(k)$	$w(k)$
0*	NA	5.99	0.018	1.28×10^{-12}	6.4×10^{-14}
1	-0.11	4.80	0.34	-0.11	6.1×10^{-5}
4*	0.15	3.007	1.68	5.61	0.302
5	-0.36	2.67	0.58	0.54	0.054
8*	0.28	2.007	1.94	18.86	0.987
9	-2.01	1.85	0.34	0.72	0.086
20*	1.98	1.005	6.54	121.81	6.21

Table 2: Investment, expected membership and welfare with mixed strategy participation. Equilibrium k denoted by *

Table 2 shows levels of welfare and expected participation for several values of k , including all of the equilibrium values when $\varepsilon = 0.11$, marked by an asterisk. There are four equilibria, with $N = 20$ giving the highest level of welfare both for investors and non-investors. Increasing ε to greater than $0.11 + 0.15 = 0.26$ removes the $k = 4$ equilibrium, and increasing ε to greater than $0.11 + 0.28 = 0.39$ removes the $k = 8$ equilibrium. Zero investment is always an equilibrium, and $k = 20$ is an equilibrium provided that $\varepsilon \leq 0.11 + 1.98 = 2.09$.

4.3 Comparison

The choice between pure and mixed strategies at the participation stage induces two types of investment games. In both of these games, there may be multiple equilibria for sufficiently low investment costs, and in both there is likely to be a boundary equilibrium in which no country invests ($k = 0$). It is also obvious that if investment cost are sufficiently low, and if a sufficiently high level of investment can drive abatement costs below the threshold at which abatement is a dominant strategy ($c = 1$), then in both games there is a second equilibrium in which all countries invest.

The interesting comparison arises when the lowest possible abatement cost remains above 1. When agents use pure strategies at the participation game, an increase in investment cost reduces the value of k at the interior equilibrium. When there are two equilibria (the boundary equilibrium $k = 0$ and an interior equilibrium), the 0-investment equilibrium has higher membership and higher welfare for both investors and non-investors. In contrast, when agents use mixed strategies at the participation stage, for sufficiently low investment costs there is a boundary equilibrium at which *all* countries invest, in addition to interior equilibria and the boundary equilibrium at which no country invests. Increasing investment costs tends to eliminate the low interior investment equilibria. In addition, expected participation and welfare are higher at higher investment equilibria.

The two types of equilibria in the participation game therefore lead to different policy implications regarding investment. The pure strategy equilibrium, emphasized by the previous literature, leads to a game in which even if investment might occur in equilibrium, that investment would likely lower abatement and welfare. This result occurs because the pure strategy participation equilibrium equals the minimally successful IEA, which is weakly increasing in abatement costs. The mixed strategy equilibrium, which we emphasize, leads to a game in which again investment might occur in equilibrium, but here it increases abatement and welfare.

The existence of multiple Pareto-ranked non-cooperative equilibria leads to a coordination problem in both models. Whereas the previous literature implies that society does better by trying to get agents to coordinate on the no-investment equilibrium, analysis of mixed strategies leads to the more intuitive conclusion that society does better by trying to get agents to coordinate on the high-investment equilibrium.

Unless the investment cost can be reduced to a level extremely close to

the threshold $c = 1$, the implications of the game with mixed strategies – like those of the game with pure strategies – are pessimistic. In our example, expected participation in the mixed strategy when all countries invest (6.54) is only slightly higher than participation in the pure strategy when no countries invest (6). Expected welfare is much higher in the former case (122 compared to 84) due to the lower abatement cost resulting from investment. However, even the larger level is only a fraction of the potential global welfare, equal to $(20 - 1.005 - 0.11) 20 = 377.7$, so both models are pessimistic about being able to capture the benefits of cooperation. We also found examples where expected participation under mixed strategies is lower but expected welfare higher, compared to the levels under pure strategies. Thus, when costs are endogenous, determined at the investment stage, the model with mixed strategies may be less pessimistic than the model with pure strategies.

In the interest of completeness, we also examined the use of mixed strategies at the investment stage. For the example in equation (7), we found that the unique mixed strategy investment equilibrium, given that countries use pure strategies at the participation stage, involves a zero probability of investment. When countries use mixed strategies at both stages, the levels of expected membership and welfare are between the highest and the lowest equilibrium levels corresponding to the scenario (reported in Table 2) in which countries use pure strategies at the investment stage and mixed strategies at the participation stage.

5 Summary

Most previous theory about IEAs uses a participation game with a pure strategy equilibrium. These models imply that the equilibrium participation is low and that investment that reduces abatement costs also reduces the incentives to join the IEA, resulting in lower equilibrium membership. Under risk neutrality and predetermined abatement costs, the use of mixed rather than pure strategies tends to reinforce the conclusion that optimal behavior by the IEA does not solve the problem of providing a global public good. For low abatement costs, however, the use of mixed strategies overturns this conclusion.

More significantly, the participation games with mixed and with pure strategies have opposite policy implications with regard to investment. In

both of the investment games induced by the two types of strategies at the participation stage, there may be multiple Pareto-ranked (pure strategy) investment equilibria. Under pure strategies at the participation stage, society would like to coordinate on a low-investment equilibrium, and under mixed strategies at the participation game, society would like to coordinate on the high-investment equilibrium. The overall policy conclusion is that when nations use mixed strategies, the promotion of investment that lowers abatement cost can ameliorate the problem of climate change, even though it is unlikely to provide a complete solution to that problem.

There are two types of risk arising with climate negotiations. We do not know whether nations will sign a treaty that leads to meaningful action, and in the event that such a treaty does emerge, we do not know the net benefit of those actions. The first type of risk is endogenous, because it depends on the interaction amongst agents, and the second type is exogenous, because it is due to imperfect knowledge and inherent uncertainty. We showed that in the presence of endogenous risk, risk aversion increases the equilibrium participation probability. Examples show that as risk aversion approaches its upper bound, the equilibrium probability of participation approaches 1. For high levels of risk aversion, countries are likely to join the IEA. In contrast, we noted that with pure strategies, the effect of an increase in risk aversion in the presence of exogenous risk has ambiguous effects on the level of participation, depending on the specifics of the model.

A Appendices

This appendix derives equation (3), explains why exogenous risk has an ambiguous effect on equilibrium participation, examines the effect of constant relative risk aversion under mixed strategies, and proves all propositions.

Derivation of equation (3) To avoid a technical complication, we assume here that $N > f(c)$, slightly stronger than our previous assumption that $N > c$. To ease notation, we define $n = N - 1$, the number of “other” countries and $k = f - 1$, the number of “other members” needed for a country to be pivotal; because $N > f$, we have $n \geq k + 1$. Using these

definitions we have

$$\begin{aligned}
\frac{G(p)}{g(p)} &= \frac{\sum_{i=k+1}^n \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}}{\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}} \\
&= \frac{\sum_{i=k}^n \left(\frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \right) - \left(\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \right)}{\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}} \\
&= \frac{\sum_{i=k}^n \left(\frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \right)}{\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}} - 1.
\end{aligned}$$

Define

$$\sigma = \frac{\sum_{i=k}^n \left(\frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \right)}{\frac{n!}{(k)!(n-k)!} p^k (1-p)^{n-k}} = \text{hypergeom} \left([1, -n+k], [1+k], \frac{p}{-1+p} \right) \quad (9)$$

where $\text{hypergeom}(\cdot)$ is the hypergeometric function. Using the definition of the hypergeometric function, and the fact that the gamma function $\Gamma(x+1) = x!$ when x is an integer, we write⁶

$$\begin{aligned}
\sigma &= \frac{\Gamma(1+k)}{\Gamma(1)\Gamma(1+k-1)} \int_0^1 \frac{t^{1-1}(1-t)^{1+k-1-1}}{\left(1-t\frac{p}{p-1}\right)^{k-n}} dt \\
&= \frac{\Gamma(1+k)}{\Gamma(k)} \int_0^1 \frac{(1-t)^{k-1}}{\left(1-t\frac{p}{p-1}\right)^{k-n}} dt = k \int_0^1 \frac{(1-t)^{k-1}}{\left(1+t\frac{p}{1-p}\right)^{k-n}} dt \quad (10) \\
&= k \int_0^1 (1-t)^{k-1} \left(1+t\frac{p}{1-p}\right)^{n-k} dt.
\end{aligned}$$

Exogenous risk Two reasonable models of exogenous risk lead to opposite comparative statics. Boucher and Bramoulle (2009) (“B&B”) distinguish between participation games that provide public goods and those that provide

⁶To check this result we also derived equation (3) without using the hypergeometric function and instead using the identity

$$\sum_{i=k}^n \left(\frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \right) = \frac{n!}{(k-1)!(n-k)!} \int_0^p t^{k-1} (1-t)^{n-k} dt$$

from Wang (1994, page 116); details are available on request.

public bads, and claim that for the former, risk aversion weakly increases and for the latter it weakly decreases participation. This distinction is puzzling, because the outcome should not depend on whether we define the action as abatement (a public good) or emissions (a public bad). In B&B the distinction does appear to matter. Here we provide a different interpretation of their result.

We normalized marginal environmental damages (the parameter “ b ” in their setting) equal to 1 and defined c as the cost-benefit ratio to a country of unilateral abatement. There are some minor technical differences depending on whether the parameter b or c is treated as random, but these are unimportant for our purposes. Here we retain the parameter b and we treat c as the random parameter. Let e equal emissions of a member of the IEA, the maximum level of which is 1, and let abatement be $a = 1 - e$. The marginal benefit to a country of emissions is c , so the marginal opportunity cost of abatement is also c . Income for an IEA member is

$$\begin{aligned} y &= A + ce - b(me + N - m) = A - b(N - m) + (c - bm)e \\ &= A - bN + bma + c(1 - a). \end{aligned} \tag{11}$$

The first line expresses income as a function of emissions, a public bad, and the second line expresses income as a function of abatement, a public good. Of course, the comparative statics of the model are exactly the same, regardless of which decision variable we choose, so in this respect at least it does not matter whether we think of the game as involving a public good or a public bad.

With this model, we can show that risk aversion, together with uncertainty about c , weakly increases abatement (decreases emissions). The IEA’s optimal level of abatement can be positive even if the number of members is $mb < f(Ec)$; with risk neutrality, optimal abatement is 0 for $mb < f(Ec)$. Consequently, risk aversion decreases the size of the “minimally successful” IEA (one that does “something useful”); risk aversion, together with exogenous risk, therefore weakly decreases the equilibrium size of the IEA under pure strategies. B&B obtain this result in a slightly different setting.

However, a reasonable alternative model leads to the opposite conclusion. Above we treated abatement costs as an opportunity cost, equal to the lost benefit of emissions. An alternative treats abatement costs as “additional”

costs. With this alternative, an IEA member's income is

$$\begin{aligned} y &= A - b(me + N - m) - c(1 - e) \\ &= A - bN + bma - ca. \end{aligned} \tag{12}$$

The first line of equation (12) treats the action as a public bad and the second treats it as a public good, but the comparative statics are (obviously) the same in both cases.

The difference between equations (11) and (12) is that the variance of income in the former case is $(1 - a)^2 \text{var}(c)$ and in the latter case it is $a^2 \text{var}(c)$. In the former case the variance of income is decreasing in a and in the latter case the variance of income is increasing in a . Risk averse agents want to decrease the variance of income.⁷ Thus, if the correct model is equation (11), risk aversion weakly increases abatement for given m , and as a consequence risk aversion weakly decreases equilibrium participation. If the correct model is equation (12) the comparative statics are reversed. This reversal has nothing to do with whether we think of agents as providing a public good or a public bad. The reversal depends on the effect that the action has on the volatility of income.

Endogenous risk with constant relative risk aversion When countries use mixed strategies and have constant relative risk aversion, an increase in the CRRA parameter η increases the participation probability. Using the equilibrium condition given in the proof to Proposition 5, it is easy to obtain numerically the equilibrium participation probability as a function of, N , A , c and η . Figure 6 shows the equilibrium p as a function of η , given $N = 20$ and $c = 5.1$ for two values of $A = 10N$ (the left scale) and $A = 1.1N$ (the right scale). For $A = 10N$, environmental damage when no nation abates is only 10% of baseline income, so for all levels of participation, income is high and the marginal utility of income is nearly constant. In this case, risk aversion has little effect on the equilibrium participation probability. In contrast, for $A = 1.1N$, environmental damage when no nation abates is approximately 90% of baseline income, so for low levels of participation income is low and risk aversion is high. In this case, risk aversion has a significant effect on the equilibrium participation probability.

⁷If agents are sufficiently risk averse, they may unilaterally choose to abate when this action reduces the variance of income (Enderes and Ohl 2003).

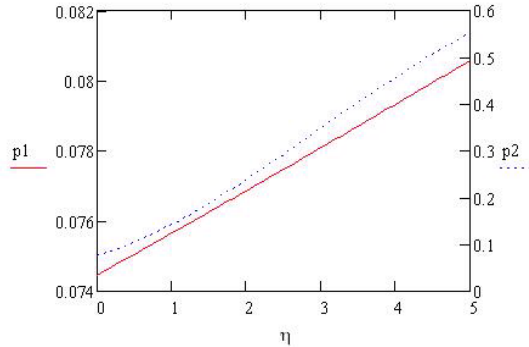


Figure 6: Relation between η and p for $N = 20$, $c = 5.1$, $A = 10N$ (left axis and the solid line $p1$) and $A = 1.1N$ (right axis and the dotted line $p2$).

In order to compare the examples with piece-wise linear and with CRRA utility, consider the case where expected income is y and actual income is $y \pm \varepsilon$, each with probability 0.5. Footnote 4 notes that for the piece-wise linear example, $w = \frac{r}{\varepsilon}$, where r is the risk premium (the amount that society would spend to stabilize income at its mean value) and ε is the size of the shock. A second order expansion of the CRRA utility function shows that

$$w = \frac{r}{\varepsilon} \approx \frac{1}{\eta} \left(-\frac{y}{\varepsilon} + \sqrt{\left(\frac{y}{\varepsilon}\right)^2 + \eta^2} \right). \quad (13)$$

The ratio $\frac{y}{\varepsilon}$ provides an inverse measure of the amount of risk relative to baseline income. As this risk becomes small ($\frac{y}{\varepsilon} \rightarrow \infty$), $w \rightarrow 0$; as the risk approaches its maximum ($\frac{y}{\varepsilon} \rightarrow 1$), $w \rightarrow \frac{1}{\eta} \left(-1 + \sqrt{1 + \eta^2} \right)$, a quantity that varies between 0 and 1, as η increases from 0 to ∞ . For example, if $\eta = 2$ (a value sometimes proposed for climate policy models) then the approximation of w given by equation (13) exceeds 0.4 provided that $\varepsilon > \frac{y}{2.1}$. Thus, moderate levels of risk aversion correspond to large values of w when the relevant risk, ε , is large.

Proof. (Proposition 1) Part (i) The fact that t is defined over a positive interval implies

$$\frac{d \left(1 + t \frac{p}{1-p} \right)}{dp} = \frac{t}{(-1+p)^2} > 0.$$

Using this expression, we have

$$\frac{d\frac{G(p)}{g(p)}}{dp} = \frac{d\sigma}{dp} = k \int_0^1 (1-t)^{k-1} (n-k) \left(1 + t \frac{p}{1-p}\right)^{n-k-1} \frac{t}{(-1+p)^2} dt > 0. \quad (14)$$

Define $x = \frac{p}{1-p}$. By using the last equality of equation 10, we have

$$\sigma(0) = \int_0^1 k(1-t)^{k-1} dt = 1,^8$$

$$\frac{d\sigma}{dx} = k \int_0^1 (1-t)^{k-1} (n-k) (1+tx)^{n-k-1} t dt > 0,$$

and

$$\frac{d^2\sigma}{dx^2} = k \int_0^1 (1-t)^{k-1} (n-k) (n-k-1) (1+tx)^{n-k-2} t^2 dt \geq 0.$$

for any $n \geq k+1$. Because σ is increasing and convex in x , the domain of x is $[0, \infty)$ and $\sigma(0) = 1$, we know that the range of σ is $[1, \infty)$. Thus $\sigma - 1$ is increasing and convex in x and the range of $\sigma - 1$ is $[0, \infty)$. When c is not an integer, $\frac{f-c}{c-1}$ is positive. Therefore, $\sigma - 1$ will cross $\frac{f-c}{c-1}$ once and only once. There is a unique solution to the equation $\frac{f-c}{c-1} = \sigma - 1$. By the one-to-one correspondence between p and x , there exists a unique p that satisfies the equilibrium condition.

Part (ii) Equation (14) shows $\frac{d\frac{G(p)}{g(p)}}{dp} > 0$. Consider an interval of c over which f is constant. Differentiating equation (2) with respect to p , treating c as a function of p we have

$$\begin{aligned} \frac{d\left(\frac{f-c}{c-1}\right)}{dc} \frac{dc}{dp} &= -\frac{f-1}{(c-1)^2} \frac{dc}{dp} = \frac{d\frac{G(p)}{g(p)}}{dp} \Rightarrow \\ \frac{dc}{dp} &= -\frac{d\frac{G(p)}{g(p)}}{dp} \frac{(c-1)^2}{f-1} < 0. \end{aligned}$$

■

The proof of Proposition 2 uses the following lemma:

⁸ $k(1-t)^{k-1}$ is the density function of $Beta(1, k)$.

Lemma 1 *In the interval $c \in (j - 1, j]$ for any integer $j \geq 3$, the expected membership in the mixed-strategy equilibrium, m , is less than j (the membership of the corresponding nontrivial pure-strategy equilibrium).*

Proof. (Lemma 1) *The first step.* When $c \in (j - 1, j]$, $f(c) = j$. The limit of the left hand side of equation (2) as c converges to $j - 1$ from the right is:

$$\lim_{c \rightarrow (j-1)^+} \frac{f - c}{c - 1} = \frac{1}{j - 2}.$$

When $j = 2$, $\lim_{c \rightarrow 1^+} \frac{f-c}{c-1} = \frac{1}{2-2} = \infty$; when $j \geq 3$, $\lim_{c \rightarrow (j-1)^+} \frac{f-c}{c-1} = \frac{1}{j-2} \leq 1$. Since

the left hand side of equation (2) decreases in c , the limit $\frac{1}{j-2}$ is the supremum of the left hand side over the interval $c \in (j - 1, j]$. For the equality to hold, $\frac{1}{j-2}$ must also be the supremum of the right hand side of equation (2).

The second step. For any integer $j \geq 3$, the supremum of the right hand side of equation (2), $\frac{1}{j-2}$, is not more than 1. We now show the right hand side of equation (2), $\frac{G(p)}{g(p)}$, evaluated at $p = \frac{j}{N}$ exceeds 1. Given that this is true, the inequality $\frac{d \frac{G(p)}{g(p)}}{dp} > 0$ implies that the equilibrium p must be strictly less than $\frac{j}{N}$ and therefore m must be less than the level in the nontrivial pure-strategy equilibrium. To confirm that the right hand side of equation (2), evaluated at $\frac{j}{N}$ exceeds 1, we evaluate the numerator and the denominator separately.

With $f(c) = j$, the first term in the sum in the numerator on the right hand side of equation (2) evaluated at $p = \frac{j}{N}$ equals

$$\begin{aligned} &= \frac{(N-1)!}{j!(N-1-j)!} \left(\frac{j}{N}\right)^j \left(1 - \frac{j}{N}\right)^{N-1-j} \\ &= \frac{(N-1)!}{(j-1)!(N-1-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-1-j} \left(\frac{j}{N} \frac{1}{j}\right) \\ &= \frac{(N-1)!}{(j-1)!(N-1-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-1-j} \frac{1}{N}. \end{aligned}$$

The denominator on the right hand side of equation (2) equals

$$\begin{aligned}
& \frac{(N-1)!}{(j-1)!(N-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-j} \\
= & \frac{(N-1)!}{(j-1)!(N-1-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-j-1} \frac{1}{N-j} \left(1 - \frac{j}{N}\right) \\
= & \frac{(N-1)!}{(j-1)!(N-1-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-j-1} \frac{1}{N}.
\end{aligned}$$

These two intermediate results imply

$$\begin{aligned}
& \frac{(N-1)!}{j!(N-1-j)!} \left(\frac{j}{N}\right)^j \left(1 - \frac{j}{N}\right)^{N-1-j} \\
= & \frac{(N-1)!}{(j-1)!(N-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-j}.
\end{aligned}$$

Thus

$$\begin{aligned}
& \sum_{i=j}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} \left(\frac{j}{N}\right)^i \left(1 - \frac{j}{N}\right)^{N-1-i} \\
= & \frac{(N-1)!}{j!(N-1-j)!} \left(\frac{j}{N}\right)^j \left(1 - \frac{j}{N}\right)^{N-1-j} \\
& + \sum_{i=j+1}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} \left(\frac{j}{N}\right)^i \left(1 - \frac{j}{N}\right)^{N-1-i} \\
= & \frac{(N-1)!}{(j-1)!(N-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-j} \\
& + \sum_{i=j+1}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} \left(\frac{j}{N}\right)^i \left(1 - \frac{j}{N}\right)^{N-1-i} \\
\geq & \frac{(N-1)!}{(j-1)!(N-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-j}.
\end{aligned}$$

Therefore, we have

$$\frac{1}{j-2} \leq 1 \leq \frac{\sum_{i=j}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} \left(\frac{j}{N}\right)^i \left(1 - \frac{j}{N}\right)^{N-1-i}}{\frac{(N-1)!}{(j-1)!(N-j)!} \left(\frac{j}{N}\right)^{j-1} \left(1 - \frac{j}{N}\right)^{N-j}}. \quad (15)$$

We conclude that the right hand side of equation (2) evaluated at $\frac{j}{N}$, given by the right hand side of inequality (15) exceeds the supremum of the left hand side of equation (2) when $j \geq 3$. Since the right hand side of equation (2) increases in p , p must be always less than $\frac{j}{N}$, and thus m is always less than j when $j \geq 3$. ■

Proof. (Proposition 2) **Part (i).** We prove Part (i) of the proposition in a series of three steps. Step 1 establishes that the functions $g(p)$ and $G(p)$ have the characteristics shown in Figure 1. Step 2 puts an upper bound and the equilibrium probability $p(c)$ and Step 3 completes the argument.

Step 1: We establish the following claims: (i) $g(0) = G(0) = 0$; (ii) $g(p)$ is single-peaked, first increasing and then decreasing in p ; (iii) $G(p)$ is increasing in p ; and (iv) $g'(0) > G'(0)$. The first claim, (i) $g(0) = G(0) = 0$ is evident by inspection. To prove claim (ii) we take the derivative of g w.r.t. p to obtain

$$\frac{dg(p)}{dp} = \frac{(N-1)!}{(f-1)!(N-f)!} p^{f-2} (1-p)^{N-f-1} [(f-1) - (N-1)p]$$

Thus when $p < \frac{f-1}{N-1}$, $\frac{dg(p)}{dp} > 0$, and when $p > \frac{f-1}{N-1}$, $\frac{dg(p)}{dp} < 0$. So $g(p)$ is single-peaked, first increasing and then decreasing in p .

Claim (iii), which states $G(p)$ is increasing in p takes a bit more work. A property of binomial distribution is that $\sum_{i=0}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i} = 1$. Thus we have

$$\begin{aligned} G(p) &= \sum_{i=f}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i} \\ &= 1 - \sum_{i=0}^{f-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i} \end{aligned} \quad (16)$$

The cumulative distribution function of a binomial distribution can be rewritten as

$$F(k; n, p) = \Pr(x \leq k) = (n-k) \frac{n!}{k!(n-k)!} \int_0^{1-p} t^{n-k-1} (1-t)^k dt$$

(Source: http://en.wikipedia.org/wiki/Binomial_distribution). Thus, we

can rewrite equation 16 as

$$\begin{aligned} G(p) &= 1 - [(N-1) - (f-1)] \frac{(N-1)!}{(f-1)!(N-f)!} \int_0^{1-p} t^{(N-1)-(f-1)-1} (1-t)^{f-1} dt \\ &= 1 - \frac{(N-1)!}{(f-1)!(N-f-1)!} \int_0^{1-p} t^{N-f-1} (1-t)^{f-1} dt \end{aligned}$$

from which we easily see that $G'(p) > 0$.

We establish claim (iv), $g'(0) > G'(0)$, using a proof by contradiction. If $g'(0) < G'(0)$, then by claims (i) (ii) and (iii), G and g will either have no intersections or have two intersections at $p > 0$, or the But Proposition 1 shows that there exists a unique (positive) equilibrium probability p . Thus it is impossible that $g'(0) < G'(0)$. The remaining possibility is that $g'(0) = G'(0)$ and that the two functions are equal in the neighborhood of $p = 0$, but that again contradicts the uniqueness of p .

Step 2: We write the equilibrium p as a function of c : $p(c)$. We now show that for any $c > 3$, $p(c) < p(c - f(c) + 3)$. Step 1 ensures that the functions $(c-1)G(p)$ and $(f-c)g(p)$ have the characteristics shown in Figure 1, and in particular that the functions $(c-1)G(p)$ and $(f-c)g(p)$ intersect at the downward sloping part of $(f-c)g(p)$. That intersection determines $p(c)$. We use the facts that

$$\begin{aligned} [f(c - f(c) + 3) - (c - f(c) + 3)] &= f(c) - c \\ (c - f(c) + 3) - 1 &< c - 1 \end{aligned}$$

(for $c > 3$). Therefore, the positive intersection of $[f(c - f(c) + 3) - (c - f(c) + 3)]g(p)$ and $[(c - f(c) + 3) - 1]G(p)$ is greater than the positive intersection of $(f - c)g(p)$ and $(c - 1)G(p)$. In other words, $p(c - f(c) + 3) > p(c)$.

Step 3: We now establish that for any $c > 2$, $p < \frac{3}{N}$. Lemma 1 establishes that for $2 < c \leq 3$, $p < \frac{3}{N}$ so we need to establish the inequality only for $c > 3$. We have . From Step 2 we know that for $c > 3$, $p(c) < p(c - f(c) + 3)$. Using the fact that $2 < c - f(c) + 3 \leq 3$ we conclude that $p(c) < p(c - f(c) + 3) < \frac{3}{N}$. Thus, we conclude that for $c > 2$, $p < \frac{3}{N}$, and consequently $m < 3$.

Part (ii) By Proposition 1, p decreases in c in the interval of $1 < c \leq 2$. So p approaches its maximum value (denoted by p_1^*) as c converges to 1 from above. Because the left hand side of equation (2) approaches infinity as c approaches 1, the right hand side of equation (2) must also approach infinity.

Thus, in the interval of $1 < c \leq 2$, as p approaches p_1^* , the right hand side of equation (2) goes to infinity. Since the right hand side of equation (2) increases in p and it is finite when $p = \frac{2}{N}$, we must have $p_1^* > \frac{2}{N}$. As p decreases with c , we must have $p > \frac{2}{N}$ when c is small enough. And when $p > \frac{2}{N}$, we have $m > 2$. Therefore, for sufficiently small c , expected membership in the mixed strategy equilibrium exceeds membership in the pure strategy equilibrium.

The right hand side of equation (2) is finite for any p not equal to 0 or 1. Because the right hand side of equation (2) approaches infinity as c approaches 1, p must approach either 0 or 1. We know that p does not approach 0 since p is decreasing over the interval and is positive inside the interval. Therefore, p approaches 1 as c approaches its lower bound 1. ■

Proof. (Proposition 3) To prove part (i) we differentiate the equilibrium condition

$$\frac{f-c}{c-1} = \sigma - 1 = k \int_0^1 (1-t)^{k-1} \left(1 + t \frac{p}{1-p}\right)^{n-k} dt - 1 \quad (17)$$

with respect to n to obtain

$$\frac{dp}{dn} = - \frac{\int_0^1 (1-t)^{k-1} \left(1 + t \frac{p}{1-p}\right)^{n-k} \ln \left(1 + t \frac{p}{1-p}\right) dt}{\int_0^1 (1-t)^{k-1} (n-k) \left(1 + t \frac{p}{1-p}\right)^{n-k-1} \frac{t}{(1-p)^2} dt} < 0.$$

Since $n = N - 1$, we have $\frac{dp}{dN} = \frac{dp}{dn} < 0$

To prove the next three parts of the proposition we substitute $m = p(n+1)$ into the last equality of 10 to obtain

$$\sigma = k \int_0^1 (1-t)^{k-1} \left(1 + t \frac{m}{n+1-m}\right)^{n-k} dt. \quad (18)$$

Differentiating m with respect to n gives

$$\frac{dm}{dn} = -\frac{A}{B} \text{ with}$$

$$A \equiv \int_0^1 (1-t)^{k-1} e^{(n-k)\ln\left(1+t\frac{m}{n+1-m}\right)} \left(\ln \left(1 + t \frac{m}{n+1-m}\right) - (n-k) \frac{\frac{tm}{(n+1-m)^2}}{1 + t \frac{m}{n+1-m}} \right) dt$$

$$B \equiv \int_0^1 (1-t)^{k-1} e^{(n-k)\ln\left(1+t\frac{m}{n+1-m}\right)} \left((n-k) \frac{\frac{t(n+1)}{(n+1-m)^2}}{1 + t \frac{m}{n+1-m}} \right) dt$$

The denominator is positive, so the sign of $\frac{dm}{dn}$ is the opposite of the sign of

$$\int_0^1 (1-t)^{k-1} e^{(n-k)\ln(1+t\frac{m}{n+1-m})} \left(\ln \left(1 + t \frac{m}{n+1-m} \right) - (n-k) \frac{\frac{tm}{(n+1-m)^2}}{1+t\frac{m}{n+1-m}} \right) dt,$$

or equivalently

$$\int_0^1 (1-t)^{k-1} e^{(n-k)\ln(1+t\frac{p}{1-p})} \left(\ln \left(1 + t \frac{p}{1-p} \right) - \frac{n-k}{n+1} \frac{tp}{(1-p)^2 \left(1 + t \frac{p}{1-p} \right)} \right) dt.$$

Now define $s(p, t) = \ln \left(1 + t \frac{p}{1-p} \right)$, $r(t) = \frac{tp}{(1-p)^2 \left(1 + t \frac{p}{1-p} \right)}$. Since $\frac{n-k}{n+1} = 1 - \frac{f}{N}$, a sufficient condition for $\frac{dm}{dN} < 0$ is, for any $t \in (0, 1)$, $\frac{s(p,t)}{r(p,t)} > 1 - \frac{f}{N}$. A sufficient condition for $\frac{dm}{dN} > 0$ is, for any $t \in (0, 1)$, $\frac{s(p,t)}{r(p,t)} < 1 - \frac{f}{N}$.

To prove part (ii), define $z = t \frac{p}{1-p}$. We have

$$\frac{s(p, t)}{r(p, t)} = (1-p) \frac{\ln(1+z)}{z} \frac{1}{1+z}. \quad (19)$$

Each term in the right hand side of 19 is positive and less than 1, so $0 < \frac{s(p,t)}{r(p,t)} < 1$. Also we know that $1 - \frac{f}{N}$ is increasing in N , $\lim_{N \rightarrow f} 1 - \frac{f}{N} = 0$, and $\lim_{N \rightarrow \infty} 1 - \frac{f}{N} = 1$. Therefore, for sufficiently large N , we have $\frac{s(p,t)}{r(p,t)} < 1 - \frac{f}{N}$ for any t , which is sufficient for $\frac{dm}{dN} > 0$; and for sufficiently small N we have $\frac{s(p,t)}{r(p,t)} > 1 - \frac{f}{N}$ for any t , which is sufficient for $\frac{dm}{dN} < 0$.

To prove part (iii) we use Proposition 2, which shows that p converges to 1 as c converges to 1. Because $\lim_{p \rightarrow 1} s = \lim_{p \rightarrow 1} r = \infty$, we can apply the L'Hopital's Rule:

$$\begin{aligned} \lim_{p \rightarrow 1} \frac{s}{r} &= \lim_{p \rightarrow 1} \frac{s_p}{r_p} \\ &= \lim_{p \rightarrow 1} \frac{\frac{t}{(1-p)^2}}{1+t\frac{p}{1-p}} \\ &= \lim_{p \rightarrow 1} \frac{t(1-p)(1-p+tp)+tp(2(1-p+tp)-t)}{(1-p)^2(1-p+tp)^2} \\ &= \lim_{p \rightarrow 1} \frac{(1-p)(1-p+tp)}{(1-p)(1-p+tp)+p(2(1-p+tp)-t)} \\ &= \frac{0}{0+t} = 0 < 1 - \frac{f}{N}. \end{aligned}$$

Therefore when c converges to 1 from the right, $\frac{s}{r}$ is less than $1 - \frac{f}{N}$ for any t , which is sufficient for $\frac{dm}{dN} > 0$.

To prove part (iv) we show that $\lim_{c \rightarrow f^-} p = 0$. We have

$$\lim_{c \rightarrow f^-} \frac{G(p)}{g(p)} = \lim_{c \rightarrow f^-} \frac{f-c}{c-1} = 0 \quad (20)$$

which implies $\lim_{c \rightarrow f^-} G(p) = \lim_{c \rightarrow f^-} \sum_{i=f}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i} = 0$. We will have either $\lim_{c \rightarrow f^-} p = 0$ or $\lim_{c \rightarrow f^-} p = 1$. Note that $G(p)$ equals the probability that at least f other countries join given that each country joins with probability p , so $\lim_{c \rightarrow f^-} p = 1$ is inconsistent with $\lim_{c \rightarrow f^-} G(p) = 0$. We conclude that $\lim_{c \rightarrow f^-} p = 0$.

Second, we use $\lim_{p \rightarrow 0} s = \lim_{p \rightarrow 0} r = 0$. Again, we apply the L'Hopital's Rule.

$$\begin{aligned} \lim_{p \rightarrow 0} \frac{s}{r} &= \lim_{p \rightarrow 0} \frac{s_p}{r_p} \\ &= \lim_{p \rightarrow 0} \frac{(1-p)(1-p+tp)}{(1-p)(1-p+tp) + p(2(1-p+tp) - t)} \\ &= \frac{1}{1} = 1 > 1 - \frac{f}{N}. \end{aligned}$$

Hence, as c converges to f from the left, p converges to 0, and $\frac{s(p,t)}{r(p,t)} > 1 - \frac{f}{N}$ for any t , which is sufficient for $\frac{dm}{dN} < 0$. ■

Proof. Proposition 4 **Part (i)**. To establish that the limiting value of $p \rightarrow 0$ as $N \rightarrow \infty$ we use the equilibrium condition

$$\frac{f-c}{c-1} + 1 = \sigma = (f-1) \int_0^1 (1-t)^{f-2} \left(1 + t \frac{p}{1-p}\right)^{N-f} dt. \quad (21)$$

If p remained strictly positive as $N \rightarrow \infty$, the right side of equation (21) would approach infinity while the left side remains constant. Therefore, p must approach 0 as $N \rightarrow \infty$.

Part (ii). We prove this claim in a series of steps. Step 1 shows that $m \equiv Np$ converges to a finite constant, defined as λ , as $N \rightarrow \infty$. Step 2

provides an intermediate result used in a subsequent calculation, and Step 3 shows that λ satisfies equation (4).

Step 1 Equation (21) implies that $\left(1 + t\frac{p}{1-p}\right)^{N-f} < \infty$ as $N \rightarrow \infty$; consequently, $(N-f)\ln\left(1 + t\frac{p}{1-p}\right) < \infty$ as $N \rightarrow \infty$. By Part (i) of the proposition, $\lim_{N \rightarrow \infty} t\frac{p}{1-p} = 0$. Therefore $\ln\left(1 + t\frac{p}{1-p}\right) \sim t\frac{p}{1-p}$ as $N \rightarrow \infty$. Thus we have $(N-f)t\frac{p}{1-p} = t\frac{N-f}{N}\frac{m}{1-p} < \infty$ as $N \rightarrow \infty$, implying that the limiting value of m (as $\rightarrow \infty$) is finite. Moreover, Proposition 3ii shows that $\frac{dm}{dN} > 0$ for N sufficiently large. Thus we know that as $N \rightarrow \infty$, m increases in N but is finite, implying that $\lim_{N \rightarrow \infty} m$ must exist. We define $\lim_{N \rightarrow \infty} m = \lambda$.

Step 2 Let $n = N-1$, and $k = f-1$. We want to show that $\lim_{n \rightarrow \infty} \left(1 + t\frac{m}{n+1-m}\right)^{n-k} = e^{t\lambda}$. To this end, let $r(n) = \left(1 + \frac{tm}{n+1-m}\right)^{\frac{n+1-m}{tm}}$ and $g(n) = tm\frac{n-k}{n+1-m}$. By Step 1, $\lim_{n \rightarrow \infty} \frac{tm}{n+1-m} = 0$. Making use of the property that $\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = e$, we have $\lim_{n \rightarrow \infty} r(n) = e$ by . Also, we have

$$\lim_{n \rightarrow \infty} g(n) = t \lim_{n \rightarrow \infty} m \left(\frac{1 - \frac{k}{n}}{1 + \frac{1}{n} - p} \right) = t \left(\lim_{n \rightarrow \infty} m \right) \left(\lim_{n \rightarrow \infty} \frac{1 - \frac{k}{n}}{1 + \frac{1}{n} - p} \right) = t\lambda,$$

where the second equality makes use of the algebraic limit theorem that $\lim_{n \rightarrow c} (a(n)b(n)) = \left(\lim_{n \rightarrow c} a(n)\right) \left(\lim_{n \rightarrow c} b(n)\right)$ if both $\lim_{n \rightarrow c} a(n)$ and $\lim_{n \rightarrow c} b(n)$ exist.

We have

$$\begin{aligned} \lim_{N \rightarrow \infty} r(n)^{g(n)} &= \lim_{N \rightarrow \infty} e^{g(n)\ln r(n)} = e^{\lim_{N \rightarrow \infty} (g(n)\ln r(n))} = e^{(\lim_{N \rightarrow \infty} g(n))(\lim_{N \rightarrow \infty} \ln r(n))} \\ &= e^{(\lim_{N \rightarrow \infty} g(n))(\ln \lim_{N \rightarrow \infty} r(n))} = \left[e^{\ln \lim_{N \rightarrow \infty} r(n)} \right]^{\lim_{N \rightarrow \infty} g(n)} = \left(\lim_{N \rightarrow \infty} r(n) \right)^{\lim_{N \rightarrow \infty} g(n)}. \end{aligned} \tag{22}$$

The second inequality in equation (22) uses the property that $\lim_{n \rightarrow c} d^{r(n)} = d^{\lim_{n \rightarrow c} r(n)}$, where d is any positive real number; the third equality is by the algebraic limit theorem. The fourth equality uses the fact that $\lim_{n \rightarrow c} \ln z(n) = \ln\left(\lim_{n \rightarrow c} z(n)\right)$ for any $z(n)$. To confirm this property, define $y(n) = \ln z(n)$.

We have $\lim_{n \rightarrow c} z(n) = \lim_{n \rightarrow c} e^{y(n)} = \lim_{n \rightarrow c} e^{\ln e^{y(n)}} = e^{\lim_{n \rightarrow c} (\ln e^{y(n)})} = e^{\lim_{n \rightarrow c} (\ln z(n))}$; taking logs of each part of this chain of equalities implies $\ln\left(\lim_{n \rightarrow c} z(n)\right) = \lim_{n \rightarrow c} (\ln z(n))$.

Thus we have

$$\lim_{n \rightarrow \infty} \left(1 + t \frac{m}{n+1-m} \right)^{n-k} = \lim_{N \rightarrow \infty} [r(n)]^{g(n)} = \left(\lim_{n \rightarrow \infty} r(n) \right)^{\lim_{n \rightarrow \infty} g(n)} = e^{t\lambda}$$

where the second inequality uses the last part of equation (22).

Step 3 We now show that equation (4) implicitly defines λ , the limiting value of \bar{m} . Using equations (17) and (18) we write the equilibrium condition as

$$\frac{1}{c-1} = \int_0^1 (1-t)^{k-1} \left(1 + t \frac{m}{n+1-m} \right)^{n-k} dt. \quad (23)$$

We have

$$\begin{aligned} \frac{1}{c-1} &= \lim_{n \rightarrow \infty} \int_0^1 (1-t)^{k-1} \left(1 + t \frac{m}{n+1-m} \right)^{n-k} dt \\ &= \int_0^1 \left[\lim_{n \rightarrow \infty} (1-t)^{k-1} \left(1 + t \frac{m}{n+1-m} \right)^{n-k} \right] dt \\ &= \int_0^1 (1-t)^{k-1} e^{t\lambda} dt. \end{aligned} \quad (24)$$

The first equality comes from taking limits of both sides of equation (23); the third equality comes from step 3; the second equality, where we pass the limit operator through the integral, is nontrivial, and requires that we use the “dominated convergence theorem”. This theorem states that the second equality in equation (24) is valid if we can find a function $l(t)$ that is independent of n and integrable (that is, $\int_0^1 l(t) dt < \infty$) and “dominates” $h_n(t)$ (that is, $|h_n(t)| \leq l(t)$), where

$$h_n(t) = (1-t)^{k-1} \left(1 + t \frac{m}{n+1-m} \right)^{n-k}$$

is the integrand in the second line of equation (24). We construct the function $l(t)$ in a separate lemma, available on request, thus establishing the validity of the second equality in equation (24).

Part (iii). When N is finite, membership of the mixed strategy equilibrium follows a binomial distribution. Parts (i) and (ii) of the proposition show that $p \rightarrow 0$ and $m \rightarrow \lambda$ when $N \rightarrow \infty$. Then by Poisson Limit Theorem, the membership of the mixed strategy equilibrium follows a Poisson distribution with parameter λ , when $N \rightarrow \infty$. ■

Proof. (Proposition 5) Part i. The equilibrium condition under risk aversion is

$$\sum_{i=f}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i} [u(A-N+i) - u(A-N+i+1-c)] = g(p) [u(A-N+f-c) - u(A-N)]. \quad (25)$$

Given any function U that satisfies our assumptions, we can rescale U so that

$$U(A-N+f-c) - U(A-N) = f-c \quad (26)$$

without changing preferences. Using equation (26), the left side of equation (25) is the same as the left side of the equilibrium condition under risk neutrality, equation (1). For $i \geq f$, we have

$$A-N+i > A-N+i+1-c > A-N+f-c. \quad (27)$$

Using equation (26) and the assumption that U is strictly concave in the neighborhood of $A-N+f-c$, inequality (27) implies

$$U(A-N+i) - U(A-N+i+1-c) < c-1$$

for all $i \geq f$. Consequently,

$$\sum_{i=f}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i} [U(A-N+i) - U(A-N+i+1-c)] < G(p)(c-1).$$

Therefore, the right side of equation (25) is less than the right side of equation (1), the equilibrium condition under risk neutrality.

Referring to Figure 1, we see that with risk aversion the hump-shape curve ($g(p)(f-c)$) does not change while the monotone curve shifts downward. As a result, the equilibrium probability is higher under risk aversion.

Part ii. Let V be a concave function, strictly concave over a subset of the interval $(A-N+f-c, A)$, so that the decision-maker with preferences V is more risk averse than the decision-maker with preferences U . Minor changes in the argument of Part i establish the result. ■

Proof. (Proposition 6) Because $k_1 > k_2$, we have $c(k_1) < c(k_2)$. We consider the case of $c(k_1) < c(k_2) < 2$ and the case of $c(k_1) < 2 < c(k_2)$ separately.

(i) $c(k_1) < c(k_2) < 2$. We have $f(c(k_1)) = f(c(k_2)) = 2$. Given $f = 2$, we can write equation (8) as

$$\begin{aligned} w &= \sum_{i=2}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} p^i (1-p)^{N-1-i} i \\ &= (N-1)p - (N-1)p(1-p)^{N-2} \\ &= (N-1)p \left[1 - (1-p)^{N-2} \right], \end{aligned}$$

which increases with p . By Proposition 1, we have $p(k_2) < p(k_1)$. So $w(k_1) > w(k_2)$.

(ii) $c(k_1) < 2 < c(k_2)$. Proposition 2 shows that $p(c) < \frac{f(c)}{N}$ when $c > 2$. Since $c(0) \geq c$ for any c , we have $p < \frac{f(c(0))}{N}$ when $c > 2$. So $p(k_2) < \frac{f(c(0))}{N}$. By Proposition 1, $p(k_1) \geq \frac{f(c(0))}{N}$ since $c(k_1) \leq \hat{c}$. Thus, we have $p(k_2) < \frac{f(c(0))}{N} \leq p(k_1)$. Moreover, $2 = f(c(k_1)) \leq f(c(k_2))$ since $c(k_1) < 2 < c(k_2)$. We have

$$\begin{aligned} w(k_2) &= \sum_{i=f(c(k_2))}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} (p(k_2))^i (1-p(k_2))^{N-1-i} i \\ &< \sum_{i=f(c(k_1))}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} (p(k_2))^i (1-p(k_2))^{N-1-i} i \\ &< \sum_{i=f(c(k_1))}^{N-1} \frac{(N-1)!}{i!(N-1-i)!} (p(k_1))^i (1-p(k_1))^{N-1-i} i \\ &= w(k_1) \end{aligned}$$

The last inequality comes from the fact that w is increasing in p when $f = 2$, as shown in case (i). ■

We include the final lemma, used in the proof of Proposition 4 for the sake of completeness.

Lemma 2 *It is valid to pass the limit operator through the integral, i.e.*

$$\lim_{n \rightarrow \infty} \int_0^1 (1-t)^{k-1} \left(1 + t \frac{m}{n+1-m} \right)^{n-k} dt = \int_0^1 \left[\lim_{n \rightarrow \infty} (1-t)^{k-1} \left(1 + t \frac{m}{n+1-m} \right)^{n-k} \right] dt.$$

Proof. (Lemma 2) We rewrite the integrand $h_n(t)$ as

$$(1-t)^{k-1} \left[\left(1 + \frac{tm}{n+1-m} \right)^{\frac{n+1-m}{tm}} \right]^{\frac{tm}{n+1-m}(n-k)}. \quad (28)$$

We now construct the function $l(t)$ that satisfies the conditions of the dominated convergence theorem (integrability and dominance), using Steps (a) and (b) as follows.

Step (a): We first show that for any $x > 0$, $(1 + \frac{1}{x})^x$ increases in x , i.e. we show that the derivative of $(1 + \frac{1}{x})^x$ is positive:

$$\begin{aligned} \frac{d(1 + \frac{1}{x})^x}{dx} &= \frac{de^{x \ln(1 + \frac{1}{x})}}{dx} = e^{x \ln(1 + \frac{1}{x})} \left[\ln \left(1 + \frac{1}{x} \right) + \frac{x \left(-\frac{1}{x^2} \right)}{1 + \frac{1}{x}} \right] \\ &= e^{x \ln(1 + \frac{1}{x})} \left[\ln \left(1 + \frac{1}{x} \right) - \frac{1}{x+1} \right] > 0. \end{aligned}$$

Note that $\lim_{x \rightarrow 0} [\ln(1 + \frac{1}{x}) - \frac{1}{x+1}] = \infty$, $\lim_{x \rightarrow \infty} [\ln(1 + \frac{1}{x}) - \frac{1}{x+1}] = 0$, and

$$\frac{d[\ln(1 + \frac{1}{x}) - \frac{1}{x+1}]}{dx} = \frac{-\frac{1}{x^2}}{1 + \frac{1}{x}} + \frac{1}{(x+1)^2} = \frac{1}{x+1} \left(\frac{1}{x+1} - \frac{1}{x} \right) < 0.$$

Thus $\ln(1 + \frac{1}{x}) - \frac{1}{x+1}$ is always positive. As a result, $\frac{d(1 + \frac{1}{x})^x}{dx}$ is always positive.

Step (b): In this step, we construct the function $l(t)$. Using the definition $x = \frac{n+1-m}{tm} = \frac{1}{t} \left(-1 + \frac{1}{p} \right)$, we have $(1 + \frac{tm}{n+1-m})^{\frac{n+1-m}{tm}} = (1 + \frac{1}{x})^x$. As n increases, p decreases (by Proposition 3(i)), thus x increases, and $(1 + \frac{1}{x})^x$ increases by Step (a). Also we know that $\lim_{n \rightarrow \infty} (1 + \frac{1}{x})^x = e$. Hence, $(1 + \frac{1}{x})^x < e$, and we have

$$h_n(t) < (1-t)^{k-1} e^{t \left[\frac{m(n-k)}{n+1-m} \right]}. \quad (29)$$

By the algebraic limit theorem $\lim_{n \rightarrow \infty} \frac{m(n-k)}{n+1-m} = \left(\lim_{n \rightarrow \infty} m \right) \left(\lim_{n \rightarrow \infty} \frac{1 - \frac{k}{n}}{1 + \frac{1}{n} - p} \right) = \lambda$.

Given any positive ε_2 , we can find n^* such that $\left| \frac{m(n-k)}{n+1-m} - \lambda \right| < \varepsilon_2$ for any

$n \geq n^*$. Note that m depends on n . Thus n^* depends on ε_2 only, independent of n . Let $l(t) = (1-t)^{k-1} e^{tv}$ where $v = \max\{\max_{n < n^*} \frac{m(n-k)}{n+1-m}, \lambda + \varepsilon_2\}$, which is independent of n . Thus $l(t)$ is independent of n . We have

$$0 < h_n(t) < (1-t)^{k-1} e^{t[\frac{m(n-k)}{n+1-m}]} < l(t). \quad (30)$$

Thus $|h_n(t)| < l(t)$, i.e., $h_n(t)$ is dominated by $l(t)$. Also, since m is finite, $\frac{m(n-k)}{n+1-m}$ is finite, v is finite, and thus $l(t)$ is finite, so $l(t)$ is integrable. As a result, by dominated convergence theorem, the second equality of 24 holds.

■

References

- BAGNOLI, M., AND B. L. LIPMAN (1988): "Successful takeovers without exclusion," *Review of Financial Studies*, pp. 89–110.
- BARRETT, S. (2003): *Environment and Statecraft*. Oxford University Press.
- BARRETT, S. (2005): "The theory of international environmental agreements," in *The Handbook of Environmental Economics*, ed. by K.-G. Maler, and J. Vincent. North Holland, Amsterdam.
- (2006): "Climate Treaties and Breakthrough Technologies," *American Economic Review*, 96, 22–25.
- BENEDICK, R. (2009): *Foreword to "Negotiating Environment and Science" by Richard Smith*. Resource for the Future, Washington DC.
- BOUCHER, V., AND Y. BRAMOULLE (2009): "Providing Global Public Goods Under Uncertainty," Working Paper.
- CHWE, M. S. (1994): "Farsighted Coalitional Stability," *Journal of Economic Theory*, 63, 299–325.
- CRAWFORD, V. P., AND J. HALLER (1990): "Leaning how to cooperate: Optimal play in repeated coordination games," *Econometrica*, 58, 571–95.
- DE ZEEUW, A. (2008): "Dynamic effects on the stability of international environmental agreements," *Journal of Environmental Economics and Management*, 55(2), 163–74.
- DIAMANTOUDI, E., AND E. SARTZETAKIS (2002): "International Environmental Agreements - The Role of Foresight," UNiversity of Aarhus working paper 2002-10.
- DIXIT, A., AND M. OLSON (2000): "Does Voluntary Participation undermine the Coase Theorem," *Journal of Public Economics*, 76, 309 – 335.
- DOWNS, G., D. ROCKE, AND P. BARSOON (1995): "Is the good news about compliance good news about cooperation?," *International Organization*, 50, 379–406.

- ENDERES, A., AND C. OHL (2003): “International Environmental Cooperation with Risk Aversion,” *International Journal of Sustainable Development*, 6, 378–392.
- EYCKMANS, J. (2001): “On the farsighted stability of the Kyoto Protocol,” CLIMNEG Working Paper 40, CORE, Universite Catholique de Louvain.
- GOLOMBEK, R., AND M. HOEL (2005): “Climate policy under technology spillovers,” *Environmental and Resource Economics*, 31, 201–227.
- (2006): “Second-best climate agreements and technology policy,” *Advances in Economic Analysis and Policy*, 6.
- (2008): “Endogenous technology and tradable emission quotas,” *Resource and Energy Economics*, 30, 197–208.
- HARSANYI, J. (1973): “Games with randomly disturbed payoffs: A new rationale for mixed strategy equilibrium points,” *International Journal of Game Theory*, 2, 1–23.
- HOEL, M., AND A. DE ZEEUW (2009): “Can a focus on breakthrough technologies improve the performance of international environmental agreements,” NBER Working Paper.
- HOLMSTROM, B., AND B. NALEBUFF (1992): “To the raider goes the surplus? A reexamination of the free-rider problem,” *Journal of Economics and Management Strategy*, 1, 37–62.
- KOHNZ, S. (2006): “Ratification quotas in international agreements: An example of emission reduction,” Munich Economics Discussion Paper 2006-10.
- MAKRIS, M. (2009): “Private provision of discrete public goods,” *Games and Economic Behavior*, 67, 292–99.
- MASKIN, E. (2003): “Bargaining, coalitions and externalities,” Presidential Address to the Econometric Society.
- MCGINTY, M. (2010): “International environmental agreements as evolutionary games,” *Environmental and Resource Economics*, 45, 251–269.

- MURDOCH, J., T. SANDLER, AND W. VIJVERBERG (2003): “The participation decision versus the level of participation in an environmental treaty,” *Journal of Public Economics*, 87, 337–62.
- MYERSON, R. (1998a): “Extended Poisson Games and the Condorcet Jury Theorem,” *Games and Economic Behavior*, 25, 111–131.
- (1998b): “Population Uncertainty and Poisson Games,” *International Journal of Game Theory*, 27, 375–292.
- (2005): “A short course on Political Economics,” Lecture notes, Peking University.
- OSMANI, D., AND R. TOL (2009): “Toward Farsightedly Stable International Environmental Agreements,” *Journal of Public Economic Theory*, 11, 455–92.
- PALFREY, T., AND H. ROSENTHAL (1984): “Participation and the provision of discrete public goods: A strategic analysis,” *Journal of Public Economics*, 24, 171–93.
- RAY, D., AND R. VOHRA (2001): “Coalitional Power and Public Goods,” *Journal of Political Economy*, 109(6), 1355 – 1382.
- RUBIO, S., AND A. ULPH (2007): “An infinite-horizon model of dynamic membership of international environmental agreements,” *Journal of Environmental Economics and Management*, 54(3), 296–310.
- SANDLER, T., AND K. SARGENT (1995): “Management of transnational commons: Coordination, publicness, and treaty formation,” *Land Economics*, 71(2), 145–62.
- WAGNER, U. (2001): “The design of stable international environmental agreements: Economic theory and political economy,” *Journal of Economic Surveys*, 15(3), 377–411.
- WANG, R. (1994): *Introduction to Probability*. Peking University Press, Beijing.
- XUE, L. (1998): “Coalitional Stability under Perfect Foresight,” *Economic Theory*, 11, 603–627.