Effective Climate Agreements under Uncertainty

TODD L. CHERRY

Center for International Climate and Environmental Research Oslo (CICERO)
P.O.Box 1129, Blindern, 0318
Oslo, Norway

Department of Economics
Appalachian State University
Boone, NC 28608

DAVID M. MCEVOY* (corresponding author)

Department of Economics
Appalachian State University
Boone, NC 28608
mcevoydm@appstate.edu
828-262-6126

Effective Climate Agreements under Uncertainty

Abstract: Meaningful international cooperation on climate change requires countries to overcome a social dilemma; collectively, countries are better off reducing greenhouse-gas emissions, but individually they are better off increasing their emissions. An effective climate agreement must motivate sovereign countries to both voluntarily agree to reduce their emissions and then comply with their commitments. Moreover, this must take place under a great deal of uncertainty regarding the damages from climate change. Our existing institutions designed to manage climate change have failed. Here we use experimental methods to test a climate agreement structure that, in theory, encourages meaningful participation and compliance. One of the defining features of the agreement is that it requires members to pay refundable deposits upon ratification. Our results show that this agreement structure can be successful at significantly reducing emissions. Most importantly, the agreement remains highly effective even in the face of uncertainty regarding the damages from a changing climate. Our results suggest that making *ex ante* deposits, even relatively small ones, serves as a coordination device that allows countries to achieve meaningful climate cooperation.

JEL Codes: D70; C91; C92; F53; H40; Q54

Keywords: climate agreements; international environmental agreements; collective action; enforcement; experiments; public goods

1. Introduction

Climate change is one of the most pressing global-scale challenges.[1] Leading national science academies have formally declared that climate change is largely due to greenhouse gas (GHG) emissions from human activity and urge nations to act to reduce global GHG emissions. Though some countries have taken noteworthy steps, over twenty-five years of international efforts have failed to yield a meaningful reduction in GHG emissions. A key source of this futility can be traced back to seventeenth century peace treaties that ended the Thirty Years' War in Europe. In establishing peace, the 1648 Treaties of Westphalia created a system of sovereignty in which each state had the authority to govern its territory. Westphalian sovereignty serves as the foundation for the current system of international law under which international obligations can be imposed on a sovereign state only with its consent. While sovereignty protects nations from external interference, it in turn requires that nations manage global-scale challenges like climate change voluntarily, through the formation of international agreements. That states only submit to a climate agreement voluntary presents a challenge for architects of an agreement because an effective agreement must first motivate sovereign countries to voluntarily commit to emission reductions and then compel them to comply with their commitments.

The Kyoto Protocol, currently the only binding international climate agreement, illustrates the difficulty of constructing an effective agreement. Kyoto has not achieved meaningful reductions in global GHG emissions precisely because of its inability to motivate the world's biggest GHG emitters to commit to emissions reductions and the lack of a credible mechanism to enforce compliance (Hovi et al. 2007; Barrett 2008; Haita 2012). Members to

[1] For example, James Hansen of the Columbia University Earth Institute, recently stated that climate change "may soon constitute a tragedy of epic proportions" (Hansen 2014). And Christiana Figueres, the Executive Secretary of the UN Framework Convention on Climate Change, argued that climate change was indeed "the most daunting challenge of the 21st century" (Figueres 2014).

Kyoto that choose to violate their commitments can do so in two ways. They can, like Canada in December of 2012, simply withdraw from the treaty thus absolving them of any emissions abatement responsibilities. Alternatively, they can remain party to the agreement and exceed their emissions cap (e.g., Japan), the only penalty being a more stringent cap levied in the next commitment period. Under this design noncompliant countries can either opt out of the agreement altogether or simply push abatement responsibilities indefinitely into the future. This enforcement provision is not credible because the penalties, if they exist at all, are insufficient to deter noncompliance. Although other compliance mechanisms are utilized in existing IEAs, they too suffer from credibility concerns. For example, the strategy of imposing trade sanctions in response to noncompliance, as incorporated in the Montreal Protocol and the Convention on the International Trade in Endangered Species, requires that the compliant parties themselves incur the costs of sanctioning; thereby undermining its credibility (Chayes and Chayes 1995). Moreover, a flimsy compliance mechanism affects participation decisions as well. Barrett 2003 (pp. 360) argues "…if the negotiators had reflected on the need for enforcement and on the difficulty of devising an effective enforcement mechanism earlier in the process, they may have negotiated a different kind of treaty – one that sustained more cooperation."

Meaningful progress on climate change requires new institutions that first motivate sovereign countries to voluntary commit to action, and second provide mechanisms that enforce the commitments. To that end, we consider a new climate agreement structure that better entices voluntary participation while also offering a credible enforcement mechanism. The basic design of the agreement is straightforward and has shown early promise in previous research (Gerber and Wichardt 2009; Hovi et al. 2012; McEvoy 2013; Cherry and McEvoy 2013). Countries individually decide whether to join an agreement, in which its members commit to previously

negotiated GHG emissions reduction targets (e.g. Kyoto and Montreal Protocols). Upon ratification, each member is required to pay a deposit to a neutral, third-party financial institution. The agreement enters into force if enough countries ratify (and pay deposits) to satisfy a participation threshold. Minimum participation requirements are standard in treaties that address international environmental issues.[2] If the threshold is not satisfied, then no deposits are paid and no agreement enters into force. If an agreement enters into force then a commitment period begins during which signatories make emissions decisions. Once the commitment period has ended and compliance is monitored, the financial institution pays back the deposit (with interest) to compliant countries. Deposits are withheld in response to detected noncompliance. Thus, a noncompliant country pays a penalty, even if they withdraw from the treaty.

If deposit amounts are set higher than the benefits of violating the agreement, the mechanism, in theory, should motivate self-interested countries to comply. The financial incentive to violate the agreement depends on the cost of meeting emissions targets and the foregone benefits of further mitigating climate change. However, there is sizeable uncertainty in these values, especially regarding the environmental damages caused by GHGs (Stern 2006; Kolstad 2007; IPCC 2007; Smith et al. 2009; Barrett 2013). A recent review of 14 published studies that estimate the economic damages from a 3 °C global temperature increase report a wide range of net impacts—from a decrease in global GDP of 12.5 percent to an increase of 2.5 percent (Tol 2009). Such uncertainty raises questions about the potential for the proposed agreement structure because *ex ante* financial deposits may not align with the cost of *ex post* compliance. Therefore we examine the performance of the proposed agreement structure when deposits are insufficient and benefits are uncertain.

[2] See Barrett 2003 for a comprehensive list of participation thresholds in international agreements.

Following previous work on climate agreements (Miliniski et al. 2008; Heitzeg et al. 2011; Tavoni et al. 2011; Barrett and Dannenberg 2012; Cherry and McEvoy 2013; Vasconcelos et al. 2014), we designed a series of laboratory public-goods experiments to empirically test the proposed agreement structure. In the absence of naturally occurring data that would allow us to examine the effectiveness of alternative climate agreements, researchers turn to experimental economics for empirical analysis (Barrett and Dannenberg 2012). Although experiments necessarily oversimplify the climate problem, the use of controlled experimentation allows researchers to shed light on specific policy components that may affect the success of an agreement. Moreover, the experiments follow economic theory that also makes many (and often more) simplifying assumptions.

The experimental design considers uncertainty in environmental damages from climate change at the time of ratification, which is reduced at the time of compliance decisions (this follows Kolstad (2007)). Although experiments necessarily simplify the process, they offer a useful empirical approach to examine the performance of agreement architectures (Falk and Heckman 2009). Results show the proposed agreement structure is highly effective. Agreements sustain meaningful participation levels, even when there is uncertainty about the benefit of contributing and even when deposits are too low to justify compliance on financial grounds. And when agreements form, compliance levels stay exceptionally high across all conditions. It appears that making *ex ante* deposits, even relatively small ones, serve as a coordination device to entice participation and an enforcement mechanism that compels compliance.

2. Experimental design and theoretical predictions

Following an established literature on global environmental problems, our underlying game is an *N* country prisoners' dilemma (Carraro and Siniscalco 1993; Rubio and Ulph 2006; Kolstad 2007; Tavoni et al. 2011; Barrett and Dannenberg 2012). In our *baseline* experiment, each player was placed in a group of five and was endowed with \$14. The players decided simultaneously how many dollars, from zero to ten, to contribute to a public account. For every dollar contributed to the public account each player earned b = \$0.60 (i.e., the marginal per capita return (MPCR) is 0.60). Since the return from the public account is less than \$1, no contributions are made in a noncooperative Nash equilibrium and each player earns \$14.[3] However because Nb>\$1, collective earnings are maximized when all five players contribute ten dollars to the public account.

In the treatments that follow, players have the opportunity to form a cooperative agreement. The basic structure remains the same across five treatments. There are two stages, and in the first (*participation stage*) players simultaneously decide whether to join an agreement. An agreement enters into force only when a minimum number of players have joined. We call this the *participation threshold*. If the participation threshold is satisfied, then in stage two the agreement members commit to contributing at least 8 dollars to the public account, and they are also required to pay a financial deposit. If the threshold is not satisfied, then no agreement forms, no deposits are paid, and the players revert back to the baseline game. Both members and nonmembers (if any) make their contribution decisions in stage two (*contribution stage*). Members are refunded their deposit if they contribute at least 8 dollars, otherwise they lose it. Nonmembers, on the other hand, make no commitments or deposits and play the baseline game.

[3] Although the Nash prediction is zero contributions, decades of experimental research reveals that players, on average, contribute positively to public goods. See Ledyard (1995) and Chaudhuri (2011) for comprehensive reviews.

Our treatment variables are the size of the participation threshold (all 5 players, or 3 of 5 players) and whether the marginal benefit of public-good contributions is certain (b = \$0.60) or uncertain (b = (\$.40 or \$.80), each with a 50/50 chance). The deposit amount was set at \$4.

In the *full-certain* treatment, all five players are required to join to satisfy the participation threshold in stage one. If the threshold is satisfied, the members each pay \$4 to an escrow account (for simplicity, zero interest is paid). In stage two each member decides how many dollars to contribute to the public account knowing that contributions less than \$8 will result in the loss of the deposit. Since the benefit of violating the agreement (\$8 - \$0.60*8 = \$3.20) is less than the penalty of \$4, a payoff-maximizing individual fully complies with the agreement and contributes \$8. Moreover, since a player earns more as a member to an agreement (\$30) than without an agreement (\$14), there is a financial incentive to join. Since all players are critical for agreement formation, there is no financial incentive not to join. An agreement with full participation and full compliance can thus be sustained in a sub-game perfect Nash equilibrium (SPNE). A failed agreement with zero contributions is also a SPNE in this game (and the treatments that follow), but it is strictly payoff dominated by the one in which all players join. Thus, when all players are required for an agreement to form and the benefit to contributing to the public good is certain, theory predicts that all players join and members fully comply and each contribute \$8.

In the *partial-certain* treatment, the participation threshold is set at three players. Therefore, agreements of sizes three, four or five can form, and when participation is less than full, members and non-members co-exist. As with full participation, agreement members are better off complying with their commitments and having their deposit returned rather than contributing less than \$8. In the participation stage, members to agreements that form earn more

than the $14 they would expect without an agreement. To demonstrate this, first consider a fully compliant agreement member's payoff function

$$\pi^m = \$14 - \$8 + s*(\$0.60*8) = \$6 + \$4.80s,$$

where s denotes the number of members to an agreement. Since $\pi^m(s \geq 3) > \pi(s < 3) = \$14$, players are always better off as members to an agreement rather than see one fail. However, non-members (for agreements of size 3 or 4) earn strictly more. A non-members payoff function is

$$\pi^{nm} = \$14 + \$4.80s,$$

and clearly $\pi^m(s \geq 3) < \pi^{nm}(s \geq 3)$. The game therefore captures the familiar tension that all players would rather be a member to an agreement before seeing an agreement fail, but would prefer being a non-member to an effective agreement above all else. Because of this tension, only agreements of size three can be sustained in a SPNE. As such, when at least 3 players are required for an agreement to form and the benefit to contributing to the public good is certain, theory predicts that an agreement with three members forms, the members fully comply and contribute $8 (and earn $20.40) and the nonmembers contribute zero (and earn $28.40).

In the uncertainty treatments, players are informed that b will be either $0.40 or $0.80 each with a 0.50 likelihood. Players are uncertain about b in the participation stage but the uncertainty is fully resolved in the contribution stage. Note that the expected marginal benefit is $0.60, so uncertainty has no predicted effect on the decision to participate for a group of risk-neutral players. The predictions developed in this section assume risk neutrality. What follows is that in the *full-uncertain* treatment, in which all five players are required, each player joins the agreement and the agreement forms. However, in the contribution stage, when the marginal benefit is revealed to be low (b = $0.40), the benefit of violating the agreement ($8 - $0.40*8 = $4.8) exceeds the deposit of $4. Theory therefore predicts, when full participation is required and

the benefit to contributing to the public good is uncertain, an agreement forms. In the contribution stage, if the benefit to contributing is low then the members violate their commitments and contribute zero (earning \$10), but if the benefit to contributing is high then members comply and contribute \$8 (earning \$38).

In the *partial-uncertain* treatment, the only difference is that only three of five players are required for an agreement to form. While uncertainty again is not predicted to effect the formation of agreements, we expect that when b = \$0.40, members violate their commitments. Thus, when partial participation is required and the benefit to contributing is uncertain, theory still predicts an agreement with three members will form. But in the contribution stage, if the benefit to contributing is low, members are predicted to violate their commitments and contribute zero earning \$10 (nonmembers earn \$14), and if the benefit is high, members are expected to fully comply and contribute \$8 and earn \$25.20 (nonmembers earn \$33.20).

Finally, we consider a treatment, called *low deposit-certain*, in which the marginal benefit of public good contributions is certain b = \$0.60, but the deposit is set at d = \$2. In this treatment the benefit of violating the treatment (\$3.20) exceeds the \$2 deposit with certainty. Here the prediction is clear; no player joins an agreement in stage one, an agreement does not form and no contributions are made to the public account. A player's earnings in this treatment are expected to mirror earnings from the baseline (\$14). For this treatment we explore only the case when full participation is required.

The experiments were computerized using software specifically designed for this research. The experiments were conducted at [**insert university**] and subjects were recruited from the general undergraduate and graduate student population. Subjects entered the laboratory, were seated and given a set of instructions that were read aloud by the moderator. 20 subjects

participated in each session. The program randomly matched five players to a group and the identities of the group members remained anonymous. Subjects played 20 rounds of play, with the groups reshuffled before the start of each period (imperfect strangers design). Subjects only participated in one of the six treatments.

The first five treatments were replicated over three sessions, yielding 1,200 individual observations per treatment. The last treatment discussed, with a low initial deposit amount (*low-deposit certain*), was replicated over two sessions yielding 800 individual observations. Some additional features were added to the experimental design to aid in the data analysis. First, before the experiment began subjects participated in a risk-elicitation exercise following Dave et al. (2010). The exercise was used to generate a risk preference measure that is included as an explanatory variable in the panel regressions. The variable isolates the effect of inherent risk preference on decision making that is not conditional on the particular features in the public-good games.

Second, when making the decision whether to join the agreement subjects were required to report their expectation of how many players would join the agreement. This feature was included to better understand the decision making process in stage one, in particular for the treatments that required only a subset of players to participate. In those treatments it is possible that a player does not join the agreement because she expects that she is not critical for its formation. In those cases, defecting is consistent with standard assumptions of profit-maximizing individuals trying to free ride off agreement members.

3. Results

We find the agreement structure is highly effective, leading to high rates of participation and compliance, and dramatically more contributions to the public good than unilateral management (baseline experiment). These findings are robust to the different participation constraints, whether or not there is uncertainty about the benefits of mitigation, and whether or not ex ante deposits are large enough to materially justify participation or compliance.

Participation. *The agreement structure led to high rates of participation and agreement formation, even when the benefit of contributing to the public good was uncertain and the ex ante deposit was materially insufficient.*
Full participation.

Table 1 shows the rates of participation and agreement formation by treatment. When ratification required full participation and contribution benefits were certain, 93.8% of players joined and agreements formed 74.1% of the time. High participation persisted even when the deposit was set too low to be incentive compatible with compliance (93.0%). When the benefit of contributing to the public good was uncertain, participation remained significant but markedly lower than in the certain case (79.3% vs. 93.8%, p <0.00).[4] This lower participation rate translated into fewer agreements being ratified when benefits were uncertain relative to certain (74.1% vs. 41.7%, p < 0.00).

Partial participation. In cases that required partial participation, we observe participation rates between 55.6% and 65.5%, which corresponds closely to the theoretic prediction of 60.0%. The lower participation is expected because a partial participation agreement introduces the potential for non-members to free ride. Agreement participation and formation registered 55.6%

[4] The p-values reported in this section are from pair-wise t (z) tests of the unconditional means (proportions).

and 63.3% when benefits were certain and 65.5% and 78.8% when benefits were uncertain ($p <$ 0.00 for both participation and agreement formation comparisons).

[TABLE 1 ABOUT HERE]

To account for the panel nature of the data (i.e., players making repeated decisions over time), we estimate three linear probability models. Table 2 presents the results. The first model pools the data from all treatments in which subjects can form cooperative agreements (all treatments except the baseline). In this model we regress the decision to join the agreement on dummies for participation threshold (dummy for partial participation was omitted), uncertainty, whether the deposit was set too low to be incentive compatible with compliance, session, period and risk preference. The risk preference variable is a subject's chosen gamble (1 to 6) from the risk-elicitation exercise. Each gamble is matched with an estimate of a subject's coefficient of relative risk aversion, and a lower gamble indicates greater risk aversion.[5] Standard errors in our regressions are robust and clustered by subject. The results are in the first column of data in Table 2 (period and session effects are suppressed).

[TABLE 2 ABOUT HERE]

Consistent with the summary results in Table 1, the pooled regression model confirms that the likelihood of joining an agreement is significantly greater when all players are required to join for entry into force. However, in the pooled model, uncertainty has no effect on the decision to

[5] See Dave et al. 2010 for discussion of how the coefficient of relative risk aversion is computed and how subjects are classified.

join the agreement, nor does the fact that a deposit is set too low to motivate purely self-interested players to join an agreement. Risk preference is also insignificant in explaining the decision to join.

The remaining two columns of Table 2 report estimates from data parsed by participation threshold. When all five players were required to join the agreement, none of the variables significantly explain changes in contribution decisions. Note in particular that the conditional effect of uncertainty is insignificant when all five players are needed, and this is in contrast to the significant finding from the unconditional tests of the aggregate data. When only three players were required to join an agreement (last column in Table 2), the likelihood of joining increases when players expect they are critical. This result is consistent with subjects trying to free ride by opting out of agreements that they expect will form without their participation. We also see that, consistent with the aggregate numbers in Table 1, uncertainty increases a player's likelihood of joining in partial participation treatments.

Compliance. *The agreement structure yields high levels of compliance across all scenarios. Members complied with their commitments even when benefits to public good contributions were uncertain and when ex ante deposits were materially insufficient.*

Table 3 shows the percentage of members that comply with their commitments. Rates of compliance are reported by the minimum participation requirement and the sufficiency of the deposit.

Full participation. When agreements required full participation and the benefits to contributing were certain, as predicted, compliance rates were close to full (96.7%). The remarkable finding is observed when benefits are uncertain and compliance decisions are made after learning whether benefits are high or low. If benefits are revealed to be high, the model

predicts compliance and we observe 96.4% compliance in these cases. If benefits are low, payoff-maximizing players should not comply. However, compliance remains high—77.5%. We also observe a striking 79% compliance level in the low-deposit treatment in which zero compliance is predicted.

Partial participation. Compliance rates among members were similar when ratification only required partial membership. When benefits were certain, compliance rates registered 94.1%, which is consistent with expectations. In cases of uncertain benefits, compliance rates remained high. When benefits were revealed to be high, compliance registered 97.8%. But, even when benefits were revealed to be low and noncompliance was predicted, compliance still reached 89.4%.

The results in Table 3 paint a clear picture; requiring financial deposits motivates high levels of compliance even when the amounts are set too low (or revealed to be too low) to be incentive compatible with compliance. The findings suggest that the dollar amount of these deposits is of secondary importance; the agreements appear to serve as coordination mechanisms to increase public good contributions, regardless of whether the deposit is set high enough to be incentive compatible with compliance.

[TABLE 3 ABOUT HERE]

Contributions. *Contributions to the public good are significant in all cases when agreements form. Contributions, however, are diminished relative to unilateral management when agreements fail to form.*

Table 4 reports average public-good contribution levels over all trials by treatment, agreement formation, and agreement membership. Contributions in the baseline treatment averaged 4.17

dollars.[6] Pairwise t-tests of average contributions between all treatments relative to the baseline are significant at the 1 percent level except for agreements requiring only three members under certainty. The difference between 4.17 and 4.33 is not statistically different from zero (p = 0.287).

Full participation. Considering only agreements that form, the agreement structure generated substantially more contributions. In the case of certain benefits, contributions under an agreement were about twice the level observed with unilateral management—8.5 vs. 4.2 ($p <$ 0.00). More surprising is that contributions remained high (7.1) even when the ex ante deposit was not large enough to expect outcomes different than the baseline.

When benefits to mitigation were uncertain, agreements still generated much higher contributions than the baseline—7.9 vs. 4.2 ($p < 0.00$). This finding arose despite compliance not being justified when benefits are revealed to be low. In such cases, members are expected to contribute nothing and forfeit their deposit, but they actually contributed substantially more than in the unilateral management baseline—6.97 vs. 4.17 ($p < 0.00$).

Partial participation. When partial participation agreements form, there are members and non-members. As Table 4 report, members contributed at levels greater than observed in the baseline (8.27 certain; 8.42 uncertain). Non-members however engage in significant free-riding and make very low contributions (0.88 with certainty; 2.27 with uncertainty). As expected, the existence of non-members that can free-ride lowers the effectiveness of the partial participation agreement relative to the full participation agreement—4.3 vs. 8.5 with certainty; 5.9 vs. 7.9 with uncertainty ($p < 0.00$ for both certain and uncertainty treatments).

[6] Contributions in the baseline game started around 50 percent and deteriorated slowly ending at approximately 33 percent of the total endowment.

When agreements form, overall contribution levels are significantly greater (at the 1 percent level) than the 4.17 average contribution from the baseline. In short, if an agreement of any type forms, contributions increase dramatically from business as usual. However, when partial agreements form, contributions by non-members are significantly less than those observed in the baseline and in cases of failed agreements. Thus, we observe players opting out of joining agreements and then free riding off the contributions of the complying members.

Failure. Table 4 reports the dramatic effect a failed attempt at cooperation has on public good provision. For agreements that do not form, contributions are markedly lower (significant at the 1% level) relative to the unilateral management baseline. This result emerges whether ratification requires full or partial participation. When participants fail to cooperate in forming an agreement, they appear to respond by contributing less than when no agreement on the table (baseline).

[TABLE 4 ABOUT HERE]

The result that cooperative agreements lead to increased public good provision in the face of uncertainty deserves more attention. Recall that the uncertainty in the game is in the return to contributions to the public account (the MPCR). Fifty percent of the time (10 of 20 trials) when players reached the second stage of the game the MPCR was revealed to be the low amount of 0.40. In these cases, the payoff maximizing decision by members is to contribute zero tokens to the public account and forfeit the deposit.[7] In short, we predicted zero contributions in half of the trials under uncertainty. In contrast, the data illustrate that even when the MPCR is revealed to be

[7] When the MPCR = 0.40 a compliant member contributes 8, receives 3.20 in return and gets back the deposit of 4. Therefore, the agreement member loses 0.80 from complying.

low, members contribute amounts much higher than in the baseline. In fact, the average contribution by agreement members when the MPCR is revealed low is 6.97 and 7.85 for the five and three player participation threshold, respectively.

As with our analysis of the participation decisions, we now make use of the panel dataset to estimate a series of models that explain contribution decisions. The dependent variable is a player's contribution to the public good. We regress this variable on a vector of dummy variables that include the participation threshold (baseline is omitted), uncertainty, the interaction between uncertainty and participation threshold, whether the deposit was low, whether the MPCR was low, the interaction between MPCR and participation threshold, risk preference, session and period. Standard errors are robust and clustered at the subject level. The results are reported in Table 5 (period and session effects are suppressed).

[TABLE 5 ABOUT HERE]

The pooled model in Table 5 is estimated using data from all treatments. Consistent with the summary statistics in Table 4, agreements requiring all players to join increase contributions to the pubic good relative to the voluntary contribution treatment. On average, agreements only requiring three of five players to join do not significantly increase contributions to the public good. Introducing uncertainty into the agreement significantly increases public good contributions. Finally, we see that when the MPCR is revealed to be low, contribution levels drop significantly. However, as discussed previously, average contributions stay far above zero when the MPCR is low.

When focusing on cases in which agreements form (second column in Table 5), contributions significantly increase relative to the baseline for all scenarios. Uncertainty has a positive influence on contributions but only at the 10 percent level. Contributions to the public good decrease when deposits are set too low (*ex ante*) and when the MPCR is revealed to be low (*ex post*). The negative effect of a low MPCR on contributions is more pronounced when agreements require all members to join for entry into force compared to only a subset of players.

The third column in Table 5 reports estimates when agreements fail. It is clear that failed agreement structures, of any participation level, cause a significant reduction in public good contributions relative to having no opportunity to form an agreement. This is just further support of the finding that failed attempts at cooperation are detrimental to public good provision.

4. Conclusions

Westphalian sovereignty complicates international cooperation to address global-scale collective action problems. This is illustrated by the struggle to achieving meaningful action on climate change. Meaningful international action on climate change requires an agreement that recognizes sovereignty while still motivating both participation and compliance. The Kyoto Protocol – currently the only binding international agreement on climate change – has failed to limit emissions to "safe" levels. Kyoto's failure, at least in part, can be attributed to its lack of a credible mechanism to enforce compliance with climate commitments. Our research empirically evaluates an agreement structure that includes an enforcement provision that is credible; one that requires countries to pay financial deposits upon ratification. The mechanism proves effective at increasing international cooperation, and this is true even when the damages from climate change are uncertain and even when deposits are too low to motivate payoff-maximizing agents to

comply. Our findings suggest that the dollar amount of the initial deposits is of secondary importance. The agreements appear to serve as coordination mechanisms to increase public good contributions (e.g., emissions abatement), regardless of whether the deposit is set high enough to be incentive compatible with compliance.

The climate agreement we consider addresses the major credibility problems with how compliance is enforced under Kyoto. Signatories to Kyoto can violate their commitments by either exceeding their emissions limits or by withdrawing from the agreement altogether, and in both cases they avoid (or delay indefinitely) financial penalties. In contrast, detected noncompliance (or withdrawal) in the climate agreement considered here triggers a loss of the initial deposit. Moreover, unlike the use of trade sanctions and other reciprocal punishment strategies, the deposit mechanism does not require that the agreement members themselves pay the cost of enforcing compliance.

The mechanism is certainly not flawless. It requires countries to pay substantial sums of money at the front end of an international agreement (possibly larger than two percent of GDP) while facing great uncertainty. It also relies on a financial institution capable of holding deposits and issuing reimbursements. It should be acknowledged that the creation of a neutral institution with the power to withhold financial deposits from noncompliant parties constitutes an exercise of international governance that is remarkably rare in international agreements. However, a climate agreement that requires *ex ante* deposits is credible, effective and simple to understand. For these reasons it should be considered seriously as a potential mechanism to enforce compliance with international climate agreements.

References

Barrett, S. 2003. *Environment and Statecraft: The Strategy of Environmental Treaty-Making*. Oxford and New York: Oxford University Press.

Barrett, S. 2008. "Climate Treaties and the Imperative of Enforcement." *Oxford Review of Economic Policy* 24(2): 239-58.

Barrett, S. 2013. "Climate Treaties and Approaching Catastrophes." *Journal of Environmental Economics and Management* 66(2): 235 – 250.

Barrett S. and A. Dannenberg. 2012. "Climate Negotiations under Scientific Uncertainty." *Proceedings of the National Academy of Sciences* 109(43): 17372-17376.

Carraro, C. and D. Siniscalco. 1993. "Strategies for the International Protection of the Environment." *Journal of Public Economics* 52(3): 309-328.

Chaudhuri, A. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature." *Experimental Economics* 14(1): 47 – 83.

Chayes, A. and Chayes, A. H. 1995. *The New Sovereignty: Compliance with International Regulatory Agreements*. Cambridge, MA: Harvard University Press.

Cherry, T. and D. McEvoy. 2013. "Enforcing Compliance with Environmental Agreements in the Absence of Strong Institutions: An Experimental Analysis." *Environmental and Resource Economics* 54(1): 63-77.

Dave, C., C.C Eckel, C.A. Johnson and C. Rojas. "Eliciting Risk Preferences: When is Simple Better?" *Journal of Risk and Uncertainty* 41(3): 219-243.

Falk, A. and J.J. Heckman. 2009. "Lab Experiments and a Major Source of Knowledge in the Social Sciences." *Science* 326(5952): 535-538.

Figueres, C. “Climate Change: Building the Will for Action.” St. Paul’s Institute, St. Paul’s Cathedral, London, May 7, 2014.

Gerber, A. and P. Wichardt. 2009. “Providing Public Goods in the Absence of Strong Institutions.” *Journal of Public Economics* 93(3-4): 429-39.

Haita, C. 2012. “The State of Compliance in the Kyoto Protocol.” *International Center for Climate Governance* Reflection No. 12/2002.

Hansen, J. 2014. “World’s Greatest Crime Against Humanity and Nature” *Climate Science, Awareness and Solutions Blog*, Columbia University Earth Institute, March 10, 2014.

Heitzig, J., K. Lessmann and Y. Zou. 2011. “Self-Enforcing Strategies to Deter Free-Riding in the Climate Change Mitigation Game and Other Repeated Public Good Games.” *Proceedings of the National Academy of Sciences* 108(38): 15739-15744.

Hovi, J., C.B. Froyn and G. Bang. 2007. “Enforcing the Kyoto Protocol: Can Punitive Consequences Restore Compliance?” *Review of International Studies* 33(3): 435-49.

Hovi, J., M. Greaker, C. Hagem and B. Holtsmark. 2012. “A Credible Compliance Enforcement System for the Climate Regime.” *Climate Policy* 12: 741 – 54.

IPCC. 2007. “Climate Change 2007.” Synthesis Report

Kolstad, C.D. 2007. “Systematic Uncertainty in Self-Enforcing International Environmental Agreements.” *Journal of Environmental Economics and Management* 53(1): 68-79.

Ledyard, J. 1995. “Public Good: A Survey of Experimental Results.” In the *Handbook of Experimental Economics*, J. Kagel and A. Roth eds., 111- 194, Princeton University Press, Princeton: New Jersey.

McEvoy, D. 2013. "Enforcing Compliance with International Environmental Agreements using a Deposit-Refund System." *International Environmental Agreements: Politics, Law and Economics* 13(4): 481-496.

Milinski, M., R.D. Sommerfeld, H. Krambeck, F.A. Reed and J. Marotzke. 2008. "The Collective-Risk Social Dilemma and the Prevention of Simulated Dangerous Climate Change." *Proceedings of the National Academy of Sciences* 105(7): 2291-2294.

Rubio, S.J. and A. Ulph. 2006. "Self-Enforcing International Agreements Revisited." *Oxford Economic Papers* 58(2): 233-263.

Smith, J.B., S.H. Schneider, M. Oppenheimer, G.W. Yohe, W. Hare, M.D. Mastrandrea, A. Patwardhan, I. Butron, J.Corfee-Morlot, C.H.D. Magadza, H-M Fussel, A.B. Pittock, A. Rahman, A. Suarez and J-P van Ypersele. 2009. "Assessing Dangerous Climate Change through an Update of the Intergovernmental Panel on Climate Change (IPCC) – Reasons for Concern." *Proceedings of the National Academy of Sciences* 106(11): 4133 – 4137.

Stern, N. 2006. "Stern Review: The Economics of Climate Change." Report prepared for the HM Treasury in the U.K. Cambridge University Press, Cambridge, U.K.

Tavoni, A., A. Dannenberg, G. Kallis and A Loschel. 2011. "Inequality, Communication and the Avoidance of Disastrous Climate Change in a Public Goods Game." *Proceedings of the National Academy of Sciences* 108(29): 11825-11829.

Tol, R.S.J. 2009. "The Economic Effects of Climate Change." *Journal of Economic Perspectives* 23(2): 29-51.

Vasconcelos, V.V., F.C. Santos, J.M. Pacheco and S.A. Levin. 2014. "Climate Policies Under Wealth Inequality." *Proceedings of the National Academy of Sciences* 111(6): 2212-2216.

Table 1: Rates of agreement participation and formation by treatment

	Full Participation		Partial Participation	
	% Joined	% Formed	% Joined	% Formed
Certainty	93.8	74.1	55.6	63.3
	(0.699)	(1.264)	(1.434)	(1.391)
	[1200]	[240]	[1200]	[240]
Uncertainty	79.3	41.7	65.5	78.8
	(1.169)	(1.423)	(1.372)	(1.181)
	[1200]	[240]	[1200]	[240]
Certainty	93.0	69.4	--	--
w/ Low Deposit	(0.902)	(1.630)		
	[800]	[160]		

Note: Standard errors are in parentheses and the number of observations is in brackets.

Table 2: Panel regression results from the participation stage

	Pooled	Full Participation	Partial Participation
Constant	0.568***	0.727***	0.511***
	(0.087)	(0.083)	(.099)
Full Participation	0.317***	---	---
	(0.056)		
Uncertainty	0.058	-0.112	0.235**
	(0.057)	(0.086)	(0.113)
Low Deposit	0.007	0.032	---
	(0.046)	(0.051)	
Risk Preference	0.009	0.006	0.012
	(0.013)	(0.013)	(0.025)
Expect to be Critical	---	---	0.146***
			(0.029)
N	5600	3200	2400
Clusters	280	160	120
Wald Chi^2	184.12***	72.15***	74.83***

Notes: Robust standard errors clustered at the subject level are in parentheses. ***,**,* indicate significant coefficients at the 1, 5 and 10 percent levels, respectively.

Table 3: Percentage of compliant members by treatment and deposit amount

	Full Participation		Partial Participation	
	Compliance Rate		Compliance Rate	
	Sufficient Deposit	Insufficient Deposit	Sufficient Deposit	Insufficient Deposit
Certainty	96.74 (0.178) [890]	--	94.14 (0.235) [512]	--
Uncertainty	96.37 (0.187) [220]	77.50 (0.418) [280]	97.80 (0.147) [363]	89.36 (0.309) [329]
Certainty w/ Low Deposit	--	78.92 (0.408) [555]	--	--

Note: Standard errors are in parentheses and the number of observations is in brackets.

Table 4: Mean public good contribution by treatment, formation, and membership

	Average Contributions			
	Pooled	Agreements Form		Agreements Fail
		Members	Non-members	
Voluntary Baseline	4.17	--	--	--
	(0.103)			
	[1200]			
Full Participation				
Certainty	7.20	8.49	--	3.49
	(0.092)	(0.055)		(0.204)
	[1200]	[890]		[310]
Uncertainty	5.47	7.846	--	3.77
	(0.116)	(0.137)		(0.143)
	[1200]	[500]		[700]
Certainty w/	6.01	7.065	--	3.60
Low Deposit	(0.134)	(0.146)		(0.219)
	[800]	[555]		[700]
Partial Participation				
Certainty	4.33	8.27	0.883	1.70
	(0.118)	(0.080)	(0.135)	(0.129)
	[1200]	[512]	[248]	[440]
Uncertainty	5.47	8.415	2.27	2.53
	(0.116)	(0.078)	(0.226)	(0.206)
	[1200]	[692]	[253]	[225]

Note: Standard errors are in parentheses and number of observations is in brackets.

Table 5: Panel regression results on public good contributions

	Pooled	Agreements Formed	Agreements Failed
Constant	4.197***	4.418***	5.184***
	(0.606)	(0.625)	(0.653)
Partial Participation	0.420	1.490**	-2.016***
	(0.694)	(0.769)	(0.650)
Full Participation	2.507***	4.073***	-1.762***
	(0.460)	(0.444)	(0.693)
Uncertainty	2.293***	1.723*	2.321***
	(0.856)	(0.913)	(0.896)
Uncertainty * Min 5	-1.057	-1.121	0.237
	(1.002)	(0.986)	(1.296)
Low Deposit	-0.928	-1.426**	1.199
	(0.597)	(0.609)	(0.835)
MPCRLow	-1.697***	-1.280***	-2.476***
	(0.267)	(0.257)	(0.534)
MPCRLow * Min5	-0.275	-0.865*	-0.082
	(0.388)	(0.483)	(0.637)
Risk Preference	0.056	0.024	-0.096
	(0.107)	(0.112)	(0.119)
N	6800	4850	3150
Clusters	340	331	336
Wald Chi^2	400.83***	307.36***	241.73***

Notes: Robust standard errors clustered at the subject level are in parentheses. ***,**,* indicate significant coefficients at the 1, 5 and 10 percent levels, respectively.