The Diplomacy of Public Goods Problems: Why States with Common Goals Lie*

Brenton Kenkel†

March 6, 2015

Abstract

Uncertainty is pervasive in international public goods problems. States may know they have a goal in common yet be unsure how much each other is willing to give to the project. According to existing theories of information transmission, states with a shared goal should be able to resolve mutual uncertainty through diplomatic talk. This paper uses a formal model to show, to the contrary, that potential contributors to a public good cannot credibly reveal any information about how much they are willing to give. States have an incentive to understate their own willingness so as to sway their partners to take up more of the burden. However, communication about private information other than a state's willingness to give may be possible. The findings illustrate why common interest, while necessary, is not a sufficient condition for political actors to reveal private information through cheap talk.

*This paper is a revised version of my dissertation chapter "Communication between Allies." I am grateful to Brett Benson, Allison Carnegie, Rob Carroll, Georgy Egorov, Mark Fey, Hein Goemans, Jeff Marshall, Curt Signorino, and Alan Wiseman for their helpful comments. The remaining errors are, of course, my own.

†Assistant Professor, Department of Political Science, Vanderbilt University. Email: `brenton.kenkel@vanderbilt.edu`.

1 Introduction

How can states efficiently provide public goods or achieve joint projects despite incentives to free-ride? The peculiar features of the international system make interstate public goods provision particularly difficult. First, there is the anarchical structure of the system, with no central authority to make policy or enforce contracts. Under anarchy, all contributions are voluntary, meaning public goods are likely to be undersupplied (Olson 1965; Bergstrom, Blume and Varian 1986). A second structural hindrance to international public goods provision is the uncertainty inherent in the international system. In joint projects just as in crises, states lack information about each other's capabilities and intentions. International relations theorists have mainly analyzed uncertainty as a cause of war (Jervis 1976; Fearon 1995), but asymmetric information is also a source of inefficiency in public goods problems (Menezes, Monteiro and Temimi 2001). While the anarchical structure of international politics would take radical upheaval to change, mutual uncertainty can in principle be resolved through communication. The crucial issue is whether states have individual incentives to misrepresent what they know, undermining the usefulness of communication. This paper examines whether, and under what circumstances, communication can help states accomplish joint projects more efficiently.

Specifically, I analyze the role of *diplomatic talk* in reducing mutual uncertainty in international public goods problems, as well as other situations where states have common goals and contributions are both costly and voluntary. By diplomatic talk, I mean ordinary communication with no strings attached, where statements cannot be verified and there is no direct cost for lying.[1] I focus on this kind of communication for two reasons. First, ordinary diplomacy is a cheap and easy form of communication among states, especially in comparison to institutional arrangements that require financial resources and sustained commitments.[2] Knowing when uncertainty can be resolved through talk alone can help scholars clarify the role of for-

[1] Ramsay (2011) calls this form of communication "simple diplomacy." I use the terms "diplomatic talk," "ordinary diplomacy," and "cheap talk diplomacy" interchangeably throughout the paper.

[2] The diplomatic system, particularly the exchange of ambassadors, can be considered an institution with its own associated costs (Anderson 1993). However, diplomatic talk as defined here includes communication as simple as a phone call between a pair of state leaders or foreign ministers.

mal institutions in international cooperative problems and can help policymakers select the best venues to achieve joint goals. Second, diplomatic talk has been identified as a means of reducing the risk of war in a recent strand of the crisis bargaining literature (Sartori 2002; Kydd 2003; Kurizaki 2007; Trager 2010, 2011; Ramsay 2011). A primary conclusion of this literature is that ordinary diplomacy can effectively raise international welfare when states have common interests, a finding that harks back to the seminal work of Crawford and Sobel (1982) on cheap talk in principal-agent problems.

No theoretical consensus exists on whether (or when) ordinary diplomacy might reduce uncertainty in international public goods provision. On one hand, the aforementioned theories of crisis bargaining lead to the conclusion that diplomatic talk ought to be effective, as common interests are stronger in public goods problems than in hostile negotiations. However, these models do not explicitly consider cooperative problems, and this conclusion runs counter to an influential line of thought in the international organizations literature. According to institutionalist theories, communication alone does not suffice to reduce uncertainty, and formal organizations must step in to supply information to their constituent states (Keohane 1982, 1984). Diplomatic talk, as defined here, is a prime example of institution-free communication. If mere talk were found to be effective at resolving uncertainty in cooperative problems, the argument in favor of establishing costly and complex international institutions to disseminate information would be undermined. Conversely, if incentives to misrepresent in public goods problems are so strong that diplomacy is ineffective, the institutionalist claim would find additional support—and the proposition that common interest enables communication would be called into question.

To resolve the theoretical dilemma, this article analyzes a formal model of diplomatic talk in a public goods problem, with an eye toward the conditions under which communication improves public goods provision in equilibrium. The brunt of the findings support the institutionalist claim that communication alone cannot substantially reduce uncertainty among states with a common goal. I focus on a scenario in which states have private information about how much they are willing to give to the joint project and examine their incentives to reveal this information through cheap talk diplomacy. The main result is that under a wide set of conditions, the supply of public goods is no greater when states can talk about their willingness to contribute to the project than if they could not communicate at all. Even though states have a shared interest in providing the public good, it is always in a state's best interest to say whatever it takes to get its partners to contribute as much as possible. This incentive to misrepresent undermines information transmission through ordinary diplomacy. My findings support not just the weak claim that formal institutions have a role to play alongside diplomacy in reducing uncertainty surrounding international public goods problems, but also the stronger claim that diplomacy alone does nothing to improve states' welfare, at least with respect to private information about willingness to give. The result contrasts starkly with usual accounts of cheap talk communication, in which talk is at least somewhat effective as long as the players' interests are aligned (Crawford and Sobel 1982).

The findings have broad implications for the study of communication and cooperation in the international sphere. One implication is that our understanding of what makes diplomatic talk effective must be revised. As I have mentioned, existing theories of diplomacy claim that communication works where common interests are present. This paper provides a counterexample, as public goods provision is an archetypical example of a mixed-motivation game, and yet diplomacy is often ineffective. My findings demonstrate that the efficacy of diplomacy depends not only on common interests among states, but also on whether a state's private information affects what it wants its partners to do. In a standard public goods problem, no matter how much a state is willing to contribute, it wants its partners to give as much as they can. States thus have no incentive to be truthful, but rather to say whatever will result in their partners giving the most. Diplomatic talk could only be effective in this setting if, for example, some states actually preferred to do the bulk of the work and allow their partners to free-ride. More generally, if we want to determine when diplomacy can affect international outcomes, we must make sure that the information a country has to reveal is related to what it wants from its partner.

Another implication of the analysis concerns how international institutions can be most effective at improving public goods provision in an environment of uncertainty. The most prominent theories of international cooperation under anarchy are built on the "tit for tat" equilibrium of the iterated prisoners' dilemma, in which countries that fail to contribute to a joint project are punished (Keohane 1984; Axelrod 1984). In real-life public goods problems, however, the question often is not whether states contribute at all, but whether what they give is in line with what they are able to give. Moreover, the most efficient feasible outcome may involve the countries that value the project most (or have the lowest cost of contribution) taking on most of the burden while their partners free-ride. States will be best able to coordinate on an efficient contribution scheme if they have full knowledge of how much their partners are able to give. Therefore, states should be punished not for failing to contribute, but rather for lying about their ability to contribute. The formal analysis shows that such dishonesty can be detected *ex post*, suggesting that it may be feasible to implement such a punishment scheme.

This paper is closely related to the economics literature on communication in public goods games. Seemingly contrary to this paper's main findings, previous studies have claimed that cheap talk may be an effective tool for resolving uncertainty about how much contributors are willing to give to a public project (Agastya, Menezes and Sengupta 2007; Barbieri 2012). The difference between these studies and the present analysis is the type of public goods problem under consideration. The existing literature examines discrete public goods games, in which the public good is supplied at full value if and only if the total contribution meets or exceeds a commonly known threshold. The assumption of a discrete good is plausible in economic settings, such as a drive to raise money for a construction project, but less so in international relations—or political settings more generally. Most often in international projects, the participants do not know exactly how much effort is required to succeed (e.g., in joint military operations), or there is no strict success-failure dichotomy (e.g., in pollution reduction). To capture scenarios like these, I model the public good as continuous rather than discrete. The logic that sustains cheap-talk communication in discrete public goods projects no longer applies in the continuous setting.[3]

In the next section of the paper, I provide an informal intuition for why states cannot reveal information to each other about how much they are willing to contribute to a public good. I then formalize the argument and prove the main result under both one-sided and two-sided uncertainty. Next, I extend the model to examine the possibility of communication under sources of private information other than a country's willingness to contribute. A concluding section summarizes the findings and discusses their implications for the study of international communication and collaboration.

2 Diplomatic Talk and Public Goods

Before delving into the formal model, I lay out a hypothetical example to illustrate the intuition behind the main argument. Imagine that India and China have a joint interest in reducing global carbon emissions. All else equal, both countries want total emissions to be as low as possible. Each country can contribute to the effort by cutting its own carbon usage. Emissions reduction is a public good, as any cut by India or China works to the benefit of both. A country pays a political cost for its own carbon-cutting efforts—due to, say, members of the winning coalition disapproving of commodity price increases—but not for its partner's. So while both India and China want to see low emissions overall, India would prefer that China do the bulk of the cutting, and vice versa. This hypothetical example is a standard public goods dilemma: the two parties have a common interest in the good being provided, but each individual participant wants to free-ride off its partner as much as possible. The countries agree about the goal but differ about how to divide the labor necessary to achieve it.

[3] In a discrete public goods game, there is no benefit to fooling one's partner into giving strictly more than the threshold. Therefore, a player may be indifferent between two cheap-talk messages even though one would induce its partner to give more. Equilibria with influential communication in the discrete models depend crucially on this indifference. In the continuous setting, a contributor is always strictly better off the more its partners give, so the possibility of indifference disappears.

Now suppose that Indian leaders are unsure of exactly how politically costly it is for Chinese leaders to reduce carbon emissions. For simplicity's sake, suppose there are just two possibilities: either China is "willing" to cut carbon usage significantly due to low political costs, or it is "unwilling," only able to contribute a token amount to the joint effort. In addition, suppose the political cost of emissions abatement in India lies between these two extremes. In public goods problems like this, the participant with the lowest cost of contribution (relative to how highly it values the stakes) typically bears a disproportionate share of the burden (Olson and Zeckhauser 1966). Therefore, the best course of action for Indian leaders depends on their beliefs about the political cost of carbon-cutting in China. If India knew China were willing, it could reduce its own emissions by relatively little, knowing that China would take up most of the effort itself. Conversely, India would do most of the carbon-cutting if it knew China had high costs and would thus be unable to contribute much on its own.

In this hypothetical example, can China credibly reveal its private information through diplomatic talk? To answer this question, we must first determine how India would respond if China revealed its willingness honestly, then analyze whether it would ever be in China's best interest to lie (Farrell and Rabin 1996). We have just seen that India will cut its own carbon emissions significantly if it learns that China says it is "unwilling," but only a bit if China claims to be "willing." Does the Indian response give either type of China an incentive to lie? Yes—it is always better for China to claim to be unwilling. Each country prefers low emissions and pays a political cost only for its own abatement efforts. Therefore, regardless of how much China is actually willing to contribute, it benefits from greater emissions reductions in India. Consider what would happen if the willing type of China claimed to be unwilling, and India believed it. China could still reduce its emissions by just as much, and thus pay the same political cost, as it would have if it had honestly revealed its willingness. Because India would cut emissions by more than if China had claimed to be willing, the result is even lower global emissions at the same political cost to China. Knowing that this incentive to feign unwillingness exists, Indian leaders should not believe China's statements about its own political costs when deciding how much to cut their own emissions.[4]

This example illustrates how it might be impossible for political actors to reveal information through cheap talk even when their interests overlap. Both of the actors in this example have the same main goal—namely, to reduce global carbon emissions—but diplomatic talk is ineffective nonetheless. An incentive to misrepresent arises because both types of China want India to give as much as possible. This finding highlights the importance of a second condition for effective communication: a state's private information must affect what it wants other states to do. The requirement of variation in preferences follows from simple game-theoretic logic (see Farrell and Rabin 1996; Baliga and Morris 2002). In order for cheap talk to influence the outcome of a game, different types of a player must choose different messages. Since cheap talk by definition has no direct effect on the outcome of a game, a player will choose the message that results in the others taking the actions she most prefers. So if all types of a player have the same preferences over others' actions, then that player will always choose the same message—meaning the message is uninformative.[5] This line of logic is exactly what makes communication impossible in the current example. China wants India to cut its own carbon usage as much as possible, regardless of China's actual political costs for contributing. Since China would always prefer for India to view it as unwilling (and thus contribute more), its claims of unwillingness are not credible.

The remainder of the paper demonstrates how diplomatic effectiveness relies on there being a relationship between a country's type and its preferences over other countries' actions. The bulk of the analysis deals with incomplete information about a country's willingness to contribute. Because a country strictly prefers to free-ride as much as possible no matter its own willingness to give, an incentive to misrepresent arises, as in the example in this section.

[4]The incentive to lie in this example is not an artifact of assuming that India reduces emissions more when China is unwilling. If the situation were reversed, and India contributed more when China claimed to be willing, then the unwilling type would have an incentive to lie.

[5]A technical caveat is in order here. Identical preferences over actions do not imply identical preferences over lotteries, so influential communication may be possible despite uniform preferences in mixed-strategy equilibria or if there is multi-sided uncertainty (Seidmann 1990; Baliga and Morris 2002). In the public goods model this paper considers, there is never a mixed-strategy equilibrium, and communication under two-sided incomplete information is ruled out explicitly in Section 5.

Under alternative sources of uncertainty, however, there may be a relationship between types and preferences over others' actions, and thus diplomacy might be effective. Indeed, when I extend the model to examine such alternatives (in the paper's penultimate section), I find circumstances in which diplomatic talk might effectively reduce uncertainty.

3 The Model

This section introduces the model of public goods, uncertainty, and communication that forms the basis of the remainder of the paper.

3.1 Contribution Game

The core of the model is a public goods game, in which two countries individually choose how much effort to contribute to a joint project or goal.[6] The players are called country 1 and country 2. In the *contribution game*, each country chooses how much effort, denoted x_i, to devote to producing the public good.[7] Let $X_i = \Re_+$ denote the set of feasible contributions for country i. The public good may be any goal or accomplishment that benefits both states, and effort is any activity that increases the value of the project at a cost to the contributor. The maximal value of the public good is normalized to 1 for each country. The proportion of the good that is produced (or the proportion of its value that is realized) is $p(x_1 + x_2)$, a strictly increasing function of the total effort contributed. I assume throughout the analysis that the function $p : \Re_+ \to [0, 1]$ is continuously differentiable and strictly concave.[8] The marginal cost of effort to each country is $c_i > 0$. A country's cost of effort is inversely related to its willingness to contribute to the public good. Although the good's realized value for both countries is determined by the total contribution, each country pays only for its own effort.

[6] Although I refer to the players as countries, the results apply to any continuous public goods problem that satisfies the assumptions below, not just in the international arena.

[7] Throughout the article, generic countries are labeled i and j, with the understanding that $i \neq j$.

[8] As in Bag and Roy (2011), these properties do not need to hold on the entire domain of p. It is sufficient that they hold on $[0, \frac{1}{\underline{c}_1} + \frac{1}{\underline{c}_2}]$, where the minimal types $\underline{c}_i$ are defined as below.

A country's payoff depends on its contribution and the proportion of the good that is produced. Given its marginal cost of effort c_i, a country's utility function is

$$u_i(x_i, x_j \mid c_i) = p(x_i + x_j) - c_i x_i. \tag{3.1}$$

A country's payoff may be non-monotonic in its own effort, as higher values of x_i increase both the production of the good and the total cost to country i. However, a country always benefits from higher contributions by its partner.

A country's marginal cost of effort determines its *stand-alone contribution*—the amount it would contribute if it were the only player in the game or, equivalently, if it expected its partner to contribute $x_j = 0$. This is the amount at which the marginal benefit of contributing equals the marginal cost, defined as

$$x^*(c) = \begin{cases} (p')^{-1}(c) & c < p'(0), \\ 0 & c \geq p'(0). \end{cases} \tag{3.2}$$

As p is strictly concave and continuously differentiable, the stand-alone contribution function x^* is continuous and strictly increasing. Because of this relationship, we may identify a country's marginal cost c_i with its willingness to contribute to the project. Important implications of the above definition are that $x^*(c) < \frac{1}{c}$ for all $c > 0$ and that any contribution $x_i > x^*(c_i)$ is strictly dominated (Bag and Roy 2011).

Under complete information, the equilibrium outcome is that the more willing country gives its stand-alone contribution, while its less willing partner contributes nothing.[9] This outcome follows from the free-riding incentives inherent to public goods problems (Olson 1965). Again, imagine two countries trying to reduce carbon emissions. The country that faces greater political costs can credibly threaten to do nothing if it knows its partner will take up the slack.

[9] If the two countries have the same marginal cost of effort, a range of outcomes are possible, as stated in the proposition below.

The more willing country cannot make the same threat successfully, as it would not be satisfied with the level of cuts its partner would make on its own. Therefore, the equilibrium entails the more-willing country contributing and the less-willing country free-riding. An interesting feature of the equilibrium is that the country with lower political costs has the same payoff as if it were acting on its own, while the higher-cost country is much better off than it would be without a partner. This difference in welfare reflects the strategic advantage of being more constrained in a cooperative problem (Schelling 1960). The following proposition formally states the complete-information equilibrium.[10]

Proposition 3.1. *Suppose the marginal contribution costs c_1 and c_2 are common knowledge. Without loss of generality, let $c_1 \leq c_2$. If $c_1 < c_2$, the unique Nash equilibrium of the contribution game is $(x_1, x_2) = (x^*(c_1), 0)$. If $c_1 = c_2 = c$, then (x_1, x_2) is a Nash equilibrium of the contribution game if and only if $x_1 + x_2 = x^*(c)$.*

Knowing each other's types allows the countries to coordinate on a contribution scheme that, while not a social welfare optimum (Bergstrom, Blume and Varian 1986), is efficient in two ways. First, the amount that is provided comes at the lowest possible total cost, since only the most willing country contributes. Second, public good provision is no less than if the more willing country acted on its own. When the countries have private information, the equilibrium outcome may fall short of even these minimal efficiency conditions.

3.2 Uncertainty and Communication

A fundamental problem of international relations is that states are uncertain of each other's capabilities and resolve (Schelling 1960; Jervis 1976; Fearon 1995). This observation applies to relations between countries with common goals as well as to adversaries. The primary goal of this article is to analyze whether diplomatic talk can allow countries with common ends to resolve their mutual uncertainty and provide public goods more effectively. To this end, I

[10]The proof of this proposition, along with all other proofs, appears in the Appendix.

introduce incomplete information and cheap talk communication to the public goods model described above.

Throughout the main portion of the article, I assume that one or both of the countries have private information about c_i, the marginal cost of contribution to the public good. This can be interpreted as private information about how much a country is willing to contribute, as c_i is inversely related to a country's stand-alone contribution. To represent this asymmetric information formally, let each country's cost parameter c_i be a random variable drawn according to the cumulative distribution function F_i. I assume that each F_i is continuous,[11] has interval support, and is common knowledge; in addition, c_1 and c_2 are drawn independently from each other. Under the assumption of interval support, we may write each country's type space as $T_i = [\underline{c}_i, \bar{c}_i]$. In the case of *one-sided incomplete information about costs*, only country 1's cost of contribution is private information: $\underline{c}_1 < \bar{c}_1$ and $\underline{c}_2 = \bar{c}_2 = c_2$, meaning F_2 is a degenerate distribution on c_2. On the other hand, under *two-sided incomplete information about costs*, we have both $\underline{c}_1 < \bar{c}_1$ and $\underline{c}_2 < \bar{c}_2$. I consider sources of incomplete information other than the marginal cost of effort in Section 6.

To model diplomatic communication, I introduce a *messaging stage* that occurs after the countries learn their types but prior to the contribution game. In the messaging stage, each country selects a message m_i from its message space M_i. For ease of exposition, I assume that each country's message space is identical to its type space, so $M_i = T_i$.[12] A *messaging strategy* is a function $\mu_i : T_i \to M_i$ that specifies the message sent by each type of country i. Messages are cheap talk, as in Crawford and Sobel (1982): each type of country i may send any message in M_i, and the chosen messages have no direct effect on either country's payoff. Payoffs in the extended game are thus given by equation (3.1), the same as in the original game.

A country's message shapes its partner's beliefs, which in turn may affect the partner's contribution. Let $\lambda_j(m_i)$ denote country j's updated beliefs about c_i after receiving the message

[11] The only exception to the assumption of continuity is the case of one-sided incomplete information, in which c_2 has a degenerate distribution and F_2 is thereby a step function.

[12] In the case of one-sided incomplete information, this means country 2 only has one message available, $M_2 = T_2 = \{c_2\}$, so we may omit the analysis of its choice in the messaging stage.

m_i, where each $\lambda_j(m_i)$ is a probability measure on T_i. For any real number r, let δ_r denote the probability measure that corresponds to a degenerate distribution on r. Each country decides how much to contribute to the joint project after receiving it's partner's message and updating its beliefs. A country's *contribution strategy* is a function $s_i : T_i \times M_i \times M_j \to X_i$, where $s_i(c_i, m_i, m_j)$ denotes the contribution by country i of type c_i after sending m_i and receiving m_j from its partner. Let $\Gamma(m_1, m_2)$ denote the contribution subgame that follows a history in which country 1 sent the message m_1 and country 2 sent m_2.

To summarize, given the messaging strategies μ_i, the beliefs λ_i, and the contribution strategies s_i for each $i = 1, 2$, the sequence of play is as follows:

1. Nature privately informs each country of its type, $c_i \in T_i$.
2. The two countries simultaneously send their messages, $\mu_i(c_i) \in M_i$.
3. Each country observes its partner's message, $\mu_j(c_j) \in M_j$, and updates its beliefs about its partner's type c_j to the probability measure $\lambda_i(\mu_j(c_j))$.
4. The two countries simultaneously contribute $s_i(c_i, \mu_i(c_i), \mu_j(c_j)) \geq 0$ to the public good.
5. The game ends and payoffs are realized.

An *assessment* $\sigma = (\mu_1, \mu_2, s_1, s_2, \lambda_1, \lambda_2)$ is a tuple containing both countries' strategies and belief systems.

3.3 Solution Concept

As this is a multistage game of incomplete information with observed actions, the appropriate solution concept is perfect Bayesian equilibrium (Fudenberg and Tirole 1991). An assessment is an *equilibrium* if each country's strategy is sequentially rational given its beliefs and the other country's strategy, and beliefs are updated in accordance with Bayes' rule whenever possible. A formal definition of equilibrium appears in the Appendix.

Equilibria can be classified in terms of how much information is revealed in the messaging stage. At one end of the informational spectrum is a *babbling equilibrium*, in which no meaningful communication takes place because neither country's messaging strategy reveals anything about its type. Formally, an equilibrium σ is babbling if both μ_1 and μ_2 are constant functions, meaning each country sends the same message regardless of its type. At the other extreme is a *fully separating equilibrium*, in which each country's messaging strategy always reveals its type: $\mu_i(c_i) = c_i$ for all c_i.[13] Any equilibrium that is neither babbling nor fully separating is *partially separating*. In the Appendix, I prove that a babbling equilibrium exists, as usual in cheap talk games.

The main task of this paper is to evaluate whether diplomatic communication can help countries provide public goods more efficiently. To conclude that communication matters, it is not enough to find an equilibrium in which countries reveal information—the information they exchange must also affect the amount they contribute. Accordingly, I focus on *influential equilibria*, in which cheap talk reveals information that in turn affects decisions in the contribution game. An equilibrium is influential if it meets two conditions:

- There is a country that partially or fully reveals its type in the messaging stage: $\mu_i(c_i) \neq \mu_i(c_i')$ for some $c_i, c_i' \in T_i$.
- That country's partner contributes a different amount depending on which message it receives: there exists a set of types $\tilde{T}_j \subseteq T_j$, where $\Pr(c_j \in \tilde{T}_j) > 0$, such that $s_j(c_j, \mu_j(c_j), \mu_i(c_i)) \neq s_j(c_j, \mu_j(c_j), \mu_i(c_i'))$ for all $c_j \in \tilde{T}_j$.

An influential equilibrium is non-babbling by definition, but a non-babbling equilibrium may not be influential. For example, an equilibrium in which country 1 fully reveals its type but country 2 always contributes $x_2 = 0$ on the path of play is non-babbling and non-influential.

The major obstacle to influential communication is that each country always prefers greater contributions by its partner. Because messages are cheap talk, there is nothing to prevent a

[13] In general, any one-to-one function μ_i corresponds to a fully separating messaging strategy; I use the form given here for expository convenience.

country from saying whatever will induce its partner to give as much as possible. It is not in a country's interest to be honest if, by lying, it could guarantee that its partner would contribute more to the joint project. In other words, there cannot be an equilibrium in which:

- Some types of a country send the message m_i, while others send a different message m_i'.
- That country's partner never contributes less, and sometimes contributes strictly more, after receiving m_i' than after receiving m_i.

In such a strategy profile, the country would always be worse off sending m_i than m_i', as the latter message guarantees an equal or greater contribution by its partner. Any type that is supposed to send m_i would have an incentive to deviate, so the strategy profile cannot be an equilibrium. The following proposition states this condition formally.

Proposition 3.2. *There is no equilibrium that satisfies all of the following criteria:*

1. *There is a pair of messages $m_i, m_i' \in M_i$, with $m_i \neq m_i'$, such that $m_i = \mu_i(c_i)$ for some $c_i \in T_i$ and $m_i' = \mu_i(c_i')$ for some $c_i' \in T_i$.*
2. *For almost all $c_j \in T_j$, $s_j(c_j, \mu_j(c_j), m_i') \geq s_j(c_j, \mu_j(c_j), m_i)$.*
3. *There is a set of types $\tilde{T}_j \subseteq T_j$ such that $\Pr(c_j \in \tilde{T}_j) > 0$ and $s_j(c_j, \mu_j(c_j), m_i') > s_j(c_j, \mu_j(c_j), m_i)$ for all $c_j \in \tilde{T}_j$.*

The subsequent analysis shows that influential communication is hard to come by—a result that can be traced back to this proposition. The problem lies in the structure of a public goods dilemma, in which a country contributes less the more it expects its partner to give. It is always in a country's best interest to feign unwillingness to contribute, so as to induce its partner to give as much as possible. In other words, a country always wants to appear to have a high marginal cost of effort, regardless of whether it truly does. The main task of the next two sections of the paper is to show that any assessment with influential communication gives rise to this incentive to misrepresent and thus, by the above proposition, cannot be an equilibrium.

4 One-Sided Uncertainty

I now examine the possibility of influential communication when there is one-sided incomplete information about costs. In this setting, both countries share an interest in the provision of a public good or the success of a joint project, and one country has private information about its cost of contributing to the common goal. The analysis focuses on whether the informed country can credibly reveal information about its type through diplomatic talk. If common interests enable communication, as claimed in previous studies of diplomacy (e.g., Kydd 2003; Trager 2010), then the two countries' joint stake in the project's success should make talk effective in this setting. To the contrary, I find that there can be no influential communication in equilibrium. The informed country always prefers to say whatever would induce its partner to contribute the most to the public good, even if this means lying about its private information. Because of this incentive to misrepresent, cheap talk statements are not credible. This finding illustrates why shared interests alone are not a sufficient condition for diplomatic communication to be effective. It also gives credence to the claim that formal organizations, or some other institutional structure besides diplomatic talk, are necessary to reduce uncertainty in international public goods problems (Keohane 1984; Martin 1992).

Influential communication is impossible under one-sided uncertainty because all types of the informed country want the same thing—namely, for the other country to contribute as much as possible. The logic of the result is similar to that of the carbon-reduction example in Section 2. In an influential strategy profile, by definition, there are two messages that country 1 sends on the path of play that evoke different contributions by country 2. But, given a choice among messages, all types of country 1 strictly prefer whichever results in country 2 expending the most effort. Therefore, in any equilibrium, country 1 must always selects the same message (so the equilibrium is babbling), or country 2 must respond to every message with the same contribution. In either case, an equilibrium cannot be influential, as stated formally in the following proposition.[14]

[14]Proposition 4.1 would still hold if we relaxed the background assumptions that p is continuous and strictly

Proposition 4.1. *There is no influential equilibrium in the game with one-sided incomplete information about costs.*

Communication fails not because of a lack of common interests, but because the informed country's preferences about its partner's actions do not vary across types. In the public goods framework modeled here, a country always prefers greater contributions by its partner, regardless of its own marginal cost of effort. All types of the informed country strictly prefer to send whichever message would result in the other country contributing the most, which in turn means that message is not a credible signal. Without this uniformity of preferences across types, influential communication might be possible. For example, suppose that low-cost types of the informed country were lone rangers who would rather contribute all effort on their own than split the burden with their partner. A situation like this, in which the incentive to free-ride is not universal, could arise if a country doubted its partner's competence or if coordinating effort from multiple sources carried transaction costs. In that case, unlike in the typical public goods setting, a low-cost type wants its partner to contribute less and thus has no incentive to feign high costs. But without that kind of variation in preferences across types, a country's diplomatic statements about its willingness to support a joint project cannot be credible and influential.

If cheap talk does not work, then states must engage in costly monitoring or intelligence activities in order to learn about each other. The question that naturally arises is whether these costs are worth paying—i.e., whether public good provision or equilibrium payoffs would be better off if the informed country's type were revealed. To answer this, I compare the equilibrium of the game with uncertainty to the complete-information benchmark. If both countries' types were common knowledge, in equilibrium the more willing country would give its standalone contribution and the less willing country would give nothing (Proposition 3.1). This outcome cannot be achieved if there is private information and no communication, as the un-

concave, that F_1 is continuous, and that F_1 has interval support. As long as p is strictly increasing, no influential equilibrium exists under one-sided incomplete information about costs.

informed country does not know whether its cost of effort exceeds its partner's. At best, the total contribution is no greater than under complete information, albeit possibly distributed inefficiently (i.e., with the higher-cost country contributing part). At worst, the total contribution is strictly less than if types were common knowledge. Therefore, the inability to communicate in the face of uncertainty indeed leaves the countries worse off, as the next result states.

Proposition 4.2. *In any equilibrium of the game with one-sided incomplete information about costs,* ex post *total contributions and social welfare are always weakly less than if c_1 were common knowledge. They are strictly less with positive probability if $\underline{c}_1 < c_2 < \bar{c}_1$ and $x^*(c_2) > 0$.*

The 2011 NATO intervention in Libya illustrates the salient features of these results. The U.S. and its allies were engaged in a joint project to protect Libyan civilians and to remove Muammar Gaddafi from power. The project's success depended on military contributions by the participants, which were costly to them for both monetary and domestic political reasons, so there was an incentive to free ride. Most importantly, there was pervasive uncertainty about the extent to which each allied country was willing to participate, particularly the U.S. (Erlanger 2011). Despite initially taking charge of coalition forces, U.S. leaders stated their desire to hand over command to NATO "within days, not weeks" of the operation beginning (Bumiller and Kirkpatrick 2011)—a disclaimer of willingness in line with what Proposition 4.1 would predict. Mutual uncertainty among the NATO allies about each other resulted in a leadership vacuum at the outset of the crisis and an aircraft shortage later on (Michaels 2013), reflecting the welfare losses identified in Proposition 4.2. Even though the allies had a common purpose, diplomatic communication did not allow them to reach agreement on which countries were best able to contribute, nor did it lead to an efficient military outcome.

5 Two-Sided Uncertainty

Now suppose that both countries have private information about how costly it is to contribute to the joint project. By allowing for mutual uncertainty, the model with two-sided incomplete information is probably a more accurate representation of the international environment than the one-sided variant. Constant changes in leadership preferences, domestic political constraints, and technological capabilities, among other factors, mean that no state can be certain of exactly how committed its partners are to achieving a joint goal. But whether the international uncertainty is unilateral or mutual, the free-riding incentives that drove the results in preceding section are constant. It remains true in the two-sided model that each state, regardless of its own type, always wants its partner to contribute as much as possible. Just as before, the universal preference to free-ride gives rise to an incentive to misrepresent, which in turn renders cheap talk diplomacy ineffective.

This section characterizes two key results about the model with two-sided incomplete information. First, there is no influential equilibrium in which the countries fully reveal their types. A country will say whatever it takes to induce its partner to contribute as much as possible—which, in this instance, entails pretending to have an extremely high marginal cost of effort. Because of this incentive to misrepresent, fully honest information exchange is impossible (except when it has no impact on contributions). Second, I examine the possibility of weaker communication plans, in which countries may only partially reveal their types. After imposing some additional structure on the model to make the analysis tractable, I find that even partial revelation is impossible in equilibrium, as the incentive to feign unwillingness remains present when countries send noisy signals of their types.

I begin by ruling out a diplomatic strategy of complete honesty, in which each country reveals exactly how much it is willing to contribute. The logic of the result is an illustration of the well-known incentive problem with full revelation in public goods games (Samuelson 1954; Clarke 1971; Green and Laffont 1977). If the countries' messaging strategies are fully revealing, then in essence there is complete information in every subgame. The countries will thus make contributions according to the complete-information equilibrium, given in Proposition 3.1. In particular, whichever country has the lower marginal cost of effort will give its stand-alone contribution, and the other country will give nothing. The equilibrium behavior in the contribution subgame creates an incentive to misrepresent in the messaging stage. Country i expects country j to give 0 if country j believes $c_j > c_i$ and $x^*(c_j)$ if it believes $c_j < c_i$. Therefore, it is always in a country's interest to claim to be the highest type, so as to maximize the chance that its partner will contribute (on the possibly false belief that it is the best-equipped to do so) rather than free-ride. This incentive to feign unwillingness drives the proof of the following result.

Proposition 5.1. *There is no fully separating influential equilibrium in the game with two-sided incomplete information about costs.*

Just as this result shows that incentives to misrepresent undermine communication via diplomacy, its proof illustrates how institutions might be able to enforce honesty. In particular, if countries expected each other to be honest, lying would be detectable *ex post*—and thus could be punished—whenever it affected the outcome of the game. Imagine that country 1's marginal cost of effort is $c_1 < \bar{c}_1$ but falsely tells country 2 its type is $\bar{c}_1$, while country 2 reports its type c_2 honestly as expected. Country 2 will contribute the most it is willing, $x_2 = x^*(c_2)$, if its cost of contribution is less than what country 1 reports as its type. If country 1's actual cost is less than country 2's, it will not be satisfied with this outcome and will "top off" the total contribution to meet its own stand-alone value, revealing itself as having lied about its willingness to give. Otherwise, if country 1's actual cost exceeds c_2, it will contribute nothing, resulting in the same outcome as if it had been honest. Because dishonesty can be detected *ex post* at least some of the time, institutions need not merely punish countries for failing to meet an arbitrary contribution threshold. Instead, they may be able to facilitate information exchange by punishing dishonesty. This strategy would benefit public goods production, since countries can coordinate on the most efficient contribution scheme only if they have all information available.

Before ruling out ordinary diplomacy as a way to communicate a country's willingness to contribute to a joint project, it is important to consider a possibility that Proposition 5.1 does not cover—partial information revelation. There are many ways for a state to reveal only a portion of its private information through diplomatic channels. Ambiguity with one's partners has a long history in diplomatic practice; for example, the U.S.-led NATO command deliberately kept members states in the dark about what it would take to "go nuclear" in a potential conflict with the Soviet Union (Barrass 2009, pp. 194–195). In terms of the formal model, if some but not all types of a country send the same message, that message is a noisy signal of its type. In some cheap talk games, this kind of partial revelation can occur in equilibrium even when full revelation cannot (e.g., Crawford and Sobel 1982). I find, however, that noisy diplomatic signaling runs into the same incentive problems as full honesty in this public goods setting.

Some additional structure on the model is needed to make the analysis of noisy diplomatic signaling tractable. First, I restrict attention to *interval messaging strategies*, in which each message corresponds to an interval of types. Given a messaging strategy μ_i, let $\mathscr{C}_i(m_i)$ denote the set of types of country i that send m_i. Formally, μ_i is an interval messaging strategy if every $\mathscr{C}_i(m_i)$ is convex. An example of noisy signaling through an interval messaging strategy would be for a country to announce that its cost of contribution is "low" if c_i is below average and "high" if it is above average.[15] Babbling and fully separating messaging strategies are both special cases of interval messaging. Second, to ease the characterization of best responses, I assume the function p is quadratic.[16] With these conditions in place, I obtain the following result ruling out influential communication via noisy diplomatic signals.

Proposition 5.2. *If p is quadratic on $[0, \frac{1}{\underline{c}_1} + \frac{1}{\underline{c}_2}]$, there is no influential equilibrium in interval*

[15] This would entail a messaging strategy μ_i such that $\mu_i(c_i) = m_L$ for all $c_i < E_{F_i}[c_i]$ and $\mu_i(c_i) = m_H$ for all $c_i > E_{F_i}[c_i]$, where $m_L \neq m_H$.

[16] The quadratic functional form is convenient because it implies that a country's optimal contribution depends only on the expected value of its partner's contribution. Although the proof of Proposition 5.2 only covers the quadratic case, I suspect that the basic logic—that every type of a country gives less after receiving a "willing" signal than an "unwilling" one—would also hold under broader conditions.

messaging strategies in the game with two-sided incomplete information about costs.

Noisy diplomatic signaling fails for essentially the same reason as total honesty: it is in each country's interest to exaggerate its costs of contributing to the international public good. Every country contributes more the less it expects its partner to give, and countries expect less from a partner with a higher marginal cost of contribution. Adding noise to diplomatic signals does not overcome the basic incentive problem—each country wants its partner to take up as much of the burden as possible, and the way to do that is to signal unwillingness to contribute.

Contrary to previous findings on diplomatic communication, the common interest in public goods provision does not give countries an incentive even to reveal a bit of their private information. Diplomatic talk can be effective only if, in addition to there being common interests, a country's private information affects its preferences about its partner's actions. In the next section, I examine whether this condition might hold—and thus whether ordinary diplomatic talk might help states succeed in joint projects—under alternative sources of uncertainty.

6 Alternative Sources of Uncertainty

The preceding analysis has shown that diplomatic talk cannot help states resolve uncertainty about each other's willingness to contribute to a joint venture. This type of uncertainty, while perhaps the most prevalent in public goods problems, is not the only one. In this section, I extend the model to analyze the effectiveness of diplomacy under two alternative sources of incomplete information. The first extension looks at whether states will share intelligence about how much effort is required for the joint project to succeed. The second extension examines the possibility of communication about comparative advantages when there are multiple avenues of contribution to the project.

6.1 Intelligence Sharing

A state may have special knowledge about how much effort it will take for a joint project to succeed. The most natural example is in the realm of military cooperation: one state may have better intelligence than its allies about the capabilities of a common enemy. Intelligence sharing presents a tradeoff similar to the one that arose in the main model. States on a joint mission have the best chances of success and can most efficiently come to a division of labor if they all have the best information possible. On the other hand, it may be in a state's individual interest to exaggerate the enemy's capabilities—or, more generally, how much effort is required for a project to succeed—so as to convince its partners to contribute more than they would have if they knew the truth. The most infamous recent example of such threat exaggeration came in U.S. Secretary of State Colin Powell's 2003 speech to the U.N. Security Council, in which he claimed to have found a "sinister nexus" between Iraq and Al Qaeda but later admitted there was no "smoking-gun, concrete evidence" of the connection (Sanger 2004). In this section, I model the intelligence-sharing problem to discover under what conditions, if any, a country will honestly communicate its discoveries to its partners. I find that a country that holds extremely accurate intelligence has no incentive to lie about what it knows. However, if a country's intelligence leaves any doubt about how much effort is required to succeed, honest communication may be impossible.

To represent the intelligence-sharing problem formally, I extend the model to assume one contributor gains additional information about the joint project. For the purposes of this extension, it is convenient to interpret the public good as a joint project with a binary outcome (success or failure), where the likelihood of success increases with the total contribution of effort. Let there be a random variable $y \geq 0$ such that realized value of the project is 1 if $x_1 + x_2 \geq y$ (success) and 0 otherwise (failure). I call y the *threshold* for success.[17] In the *intelligence-sharing game*, Nature sends country 1 a signal of the threshold before contribu-

[17]On threshold uncertainty in public goods games, see Nitzan and Romano (1990), Suleiman (1997), and McBride (2006).

tions are chosen, and country 1 can tell its partner about the intelligence it has received. The threshold and the signal are drawn from the joint probability distribution $F_{y,z}$ on $Y \times Z$, where $Z \subseteq Y \subseteq \Re_+$ and F_y is the marginal distribution of y. This distribution is common knowledge, but only country 1 observes the realized value of z, and neither country observes the realized value of y. For each $z \in Z$, let $F_y(\cdot \mid z)$ denote the conditional distribution function of the threshold given the signal. The signal is *perfectly informative* if $z = y$, meaning country 1 learns the exact contribution level required for success. With a perfectly informative signal, the game resembles an ordinary discrete public goods model (at least from country 1's perspective), as the conditional distribution function is a step function,

$$F_y(y' \mid z) = \begin{cases} 0 & y' < z, \\ 1 & y' \geq z. \end{cases}$$

As before, I assume that the set of messages country 1 may send is the same as its type space, $M_1 = Z$. In order to focus on the intelligence-sharing problem, the marginal costs of contribution, c_1 and c_2, are assumed to be common knowledge in this extension.

The analysis of the main model showed that common interest is not enough to make diplomatic talk effective. For a state to maintain credibility, its private information must affect what it wants its partner to do—otherwise, the state will always prefer to say the same thing regardless of its actual type. Like the main model, the intelligence-sharing extension fails this requirement. Whether it believes the threshold for success is low or high, the informed country wants its partner to contribute as much as possible. We should therefore expect diplomatic talk to be undermined by incentives to misrepresent, as in the preceding analysis. There is, however, one subtle way in which the intelligence-sharing game differs. Suppose intelligence is perfect, so the informed country knows the exact threshold for success. The informed country then has no incentive to induce its partner to contribute more than the threshold, since doing so has no additional effect on the chance of success. Because the informed country does not benefit from over-contribution, there is not the same incentive to misrepresent as in the original model, and influential communication is indeed possible. Nevertheless, even minor relaxations of the perfect-intelligence assumption can revive the incentive to exaggerate, negating the effectiveness of diplomatic talk.

Under perfectly informative signaling, in which $z = y$, there is an equilibrium in which the informed country always tells its partner exactly what signal it received. What sustains the equilibrium, besides the lack of incentive for over-contribution mentioned above, is that the informed country is compensated for sharing its intelligence by getting to free-ride on its partner as much as possible. The contributions on the path of play take one of three forms, depending on the signal country 1 receives (and hence the true value of the threshold for success):

1. *The threshold does not exceed country 2's willingness to contribute:* $z \leq \frac{1}{c_2}$.

   Country 2 gives the full amount ($x_2 = z$), and country 1 gives nothing ($x_1 = 0$).

2. *The threshold exceeds country 2's willingness to contribute, but not the countries' total willingness:* $\frac{1}{c_2} < z \leq \frac{1}{c_1} + \frac{1}{c_2}$.

   Country 2 gives as much as it is willing ($x_2 = \frac{1}{c_2}$), and country 1 makes up the difference ($x_1 = z - \frac{1}{c_2}$).

3. *The threshold exceeds the countries' total willingness to contribute:* $z > \frac{1}{c_1} + \frac{1}{c_2}$.

   Neither country contributes ($x_1 = x_2 = 0$).

Figure 1 illustrates each country's contribution as a function of the threshold. Why is it always in country 1's interest, given these strategies, to report its signal honestly? In the first case, country 1 receives its most preferred outcome—assured success at no cost to itself—and thus has no incentive to deviate. In the second case, there is no message country 1 could send that would induce country 2 to contribute more, so again there is no incentive to deviate. The trickiest case is the third. The most country 1 could induce its partner to contribute by falsely

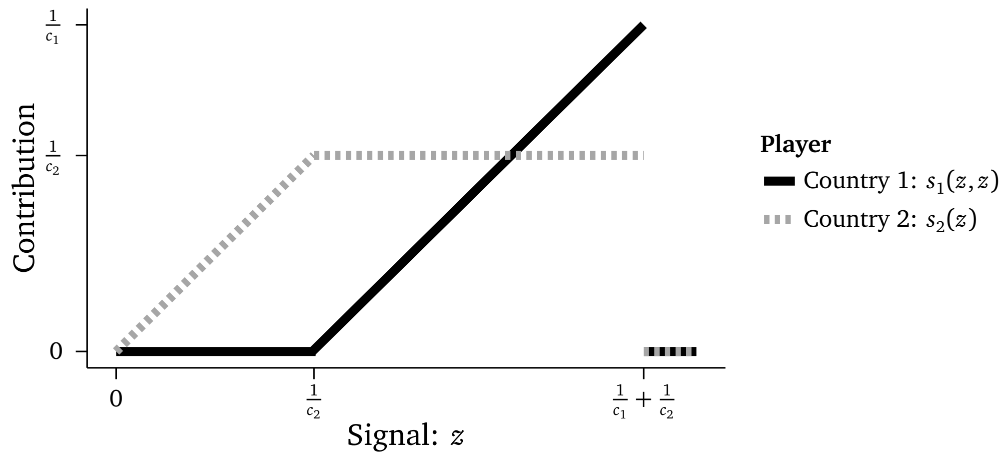

Figure 1. Contributions made by each country on the path of play in the equilibrium given in Proposition 6.1, as a function of the signal received by country 1.

reporting its signal is $x_2 = \frac{1}{c_2}$. Even then, the leftover amount required for success would exceed $\frac{1}{c_1}$, which is the most country 1 is willing to contribute. The net result of diplomatic dishonesty would be that country 1 contributes nothing and the project fails—the same as if it followed the proposed strategy. Therefore, under perfectly informative signaling, there is no barrier to honest diplomatic communication. The following proposition summarizes the result.

Proposition 6.1. *If the signal is perfectly informative, then there is a fully separating, influential equilibrium of the intelligence-sharing game.*

What makes diplomacy work under perfectly informative signaling is that the informed country has no incentive to induce its partner to over-contribute. The incentive to lie reappears if the informed country cannot place bounds on exactly how much effort is required for success. If the informed country always believes that additional contributions raise the chance of success, then it always strictly prefers the greatest possible contribution by its partner, meaning the same incentive to misrepresent arises as in the original model. Diplomatic talk is thus an ineffective vehicle for intelligence sharing if the conditional distribution of the threshold always has full support. Adding a small amount of noise can trigger this condition—for example, consider an "almost perfectly informative" signal that equals the true threshold y with probability $1-\epsilon$ and is drawn from $U[0, \frac{1}{c_1} + \frac{1}{c_2}]$ with probability ϵ (where $0 < \epsilon < 1$). The following proposition states the no-communication result.

Proposition 6.2. *If $[0, \frac{1}{c_1} + \frac{1}{c_2}] \subseteq \operatorname{supp} F_y(\cdot \,|\, z)$ for all signals z, then there is no influential equilibrium of the intelligence-sharing game.*

Two major conclusions emerge from this analysis. First, states need channels besides ordinary diplomacy to reliably share information that bears on the success of joint projects. We have seen that diplomacy is least likely to be effective for intelligence sharing in situations where uncertainty is greatest, which in turn are the most likely to be dangerous for the states involved (Jervis 1976; Fearon 1995). One way for countries with common goals to overcome incentive problems is to directly integrate their information-gathering and -dissemination systems, as NATO members did in establishing an Intelligence Fusion Center in 2007 (Mitchell 2007). Second, the results reinforce the claim that common interest alone is insufficient for diplomatic talk to be an effective means of information revelation. Diplomacy cannot work unless a country's private information has a nontrivial effect on what it wants its partner to do. In the present context, that condition is met—and thus influential communication through mere talk is possible—only when intelligence is of implausibly high quality.

6.2 Multiple Contribution Methods

So far I have assumed that all contributions to the joint project are the same. In real-life public goods problems, however, there are often multiple ways to contribute. Examples abound: The fight against global warming depends on reductions in various greenhouse gases, including both carbon dioxide and methane. Global public health programs require the manufacture of supplies (e.g., vaccinations), distribution of those supplies to affected populations, and education and outreach efforts. Contemporary military interventions usually require both ground troops and air support. Not only may states differ in their willingness or ability to supply various methods of contribution, but they may also be uncertain about each other's relative advantages. In this section, I modify the model to allow for two means of contribution and cross-country differences in the cost of each method. Unlike in the preceding analyses, I find that diplomatic communication is effective in this setting.

When there is more than one way to contribute, a country does not necessarily always want the same thing from its partner—and that is precisely why diplomacy works here. To see why, consider the example of a military venture that requires ground forces and air power to succeed. A country with an excellent infantry may prefer for its partner to focus on the air war, whereas a country that specializes in air power may want its partner to supply the ground troops. Therefore, a country with private information about its specialty might have an incentive to be honest with its partner, so as to induce it to cooperate the right way. The finding is similar to Trager's (2011) results about multidimensional bargaining between adversaries.

The formal definition of the *two-method game* is as follows. In the contribution stage, each state decides whether to devote effort to either or both of two avenues of contribution. To keep the analysis simple, I assume the contribution to each method is dichotomous.[18] Formally, each state's action in the contribution stage is a pair $x_i = (x_i^A, x_i^B)$ of binary decisions, denoting the country's decision to contribute to methods A and B respectively. Each country's action space is $X_i = \{0,1\}^2$. The cost to country i of participating in method A is α_i, and its cost of method B is β_i. The probability of success is now a function of the total contribution to each method, denoted $p(x^A, x^B)$, which is strictly increasing in both arguments. A country's utility function in the two-method contribution game is

$$u_i(x_i, x_j \mid \alpha_i, \beta_i) = p(x_i^A + x_j^A, x_i^B + x_j^B) - \alpha_i x_i^A - \beta_i x_i^B. \tag{6.1}$$

Figure 2 contains the payoff matrix for the contribution game.

I assume the source of uncertainty is country 1's cost of contributing to method B. Country 1's type β_1 equals β_1^L with probability $\pi \in (0,1)$ and β^H with probability $1-\pi$, where

[18] The main finding here—that an influential equilibrium exists in a nontrivial subset of the parameter space—would still hold if the action space were continuous.

		Country 2			
		$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
Country 1	$(0,0)$	$p(0,0)$ $p(0,0)$	$p(1,0)$ $p(1,0)-\alpha_2$	$p(0,1)$ $p(0,1)-\beta_2$	$p(1,1)$ $p(1,1)-\alpha_2-\beta_2$
	$(1,0)$	$p(1,0)-\alpha_1$ $p(1,0)$	$p(2,0)-\alpha_1$ $p(2,0)-\alpha_2$	$p(1,1)-\alpha_1$ $p(1,1)-\beta_2$	$p(2,1)-\alpha_1$ $p(2,1)-\alpha_2-\beta_2$
	$(0,1)$	$p(0,1)-\beta_1$ $p(0,1)$	$p(1,1)-\beta_1$ $p(1,1)-\alpha_2$	$p(0,2)-\beta_1$ $p(0,2)-\beta_2$	$p(1,2)-\beta_1$ $p(1,2)-\alpha_2-\beta_2$
	$(1,1)$	$p(1,1)-\alpha_1-\beta_1$ $p(1,1)$	$p(2,1)-\alpha_1-\beta_1$ $p(2,1)-\alpha_2$	$p(1,2)-\alpha_1-\beta_1$ $p(1,2)-\beta_2$	$p(2,2)-\alpha_1-\beta_1$ $p(2,2)-\alpha_2-\beta_2$

Figure 2. Payoff matrix for the contribution stage of the two-method game.

$\beta_1^L < \alpha_1 < \beta_1^H$. I call country 1 "$B$-advantaged" if $\beta_1 = \beta_1^L$ and "A-advantaged" if $\beta_1 = \beta_1^H$. The prior distribution of β_1 is common knowledge, but only country 1 observes its realized value. The other cost parameters, α_1, α_2, and β_2, are common knowledge. I assume it is never in a state's interest to contribute to a method that its partner is also contributing to.[19] This condition is equivalent to

$$\begin{aligned} p(2,x^B) - p(1,x^B) &\leq \min\{\alpha_1,\alpha_2\} \quad \text{for all } x^B = 0,1,2, \\ p(x^A,2) - p(x^A,1) &\leq \min\{\beta_1^L,\beta_2\} \quad \text{for all } x^A = 0,1,2. \end{aligned}$$

Between when β_1 is realized and when contributions are chosen, country 1 may send a message about its type to its partner. As before, the messaging space and type space are identical: $M_1 = T_1 = \{\beta_1^L, \beta_1^H\}$.

I look for a variety of equilibrium in which the two countries coordinate on a natural division of labor: country 1 contributes to whichever method it is advantaged in, and country 2 takes up the other. Reaching such a division of labor requires that country 1 honestly reveal its type. A *division of labor equilibrium* is thus an equilibrium in which:

1. Country 1 honestly reveals its type in the messaging stage.

2. If country 1 is A-advantaged, country 1 contributes $(x_1^A, x_1^B) = (1,0)$ and country 2 con-

[19] Influential equilibria would still exist if this condition were relaxed, though the statement of the conditions for Proposition 6.3 would become more cumbersome.

tributes $(x_2^A, x_2^B) = (0, 1)$.

3. If country 1 is B-advantaged, country 1 contributes $(x_1^A, x_1^B) = (0, 1)$ and country 2 contributes $(x_2^A, x_2^B) = (1, 0)$.

In a division of labor equilibrium, the total contribution along the path of play is always $(x^A, x^B) = (1, 1)$. A division of labor equilibrium is influential by definition, though other types of influential equilibria may also exist.[20]

In order to delineate the conditions under which there is a division of labor equilibrium, it is useful to think of country 1's messages as statements about what it will do in the contribution stage. Consider the B-advantaged type's message, which can be interpreted as the statement "I will contribute to method B." Farrell and Rabin (1996) delineate two requirements for a statement like this to be credible, namely that it be *self-committing* and *self-signaling*. A statement is self-committing if the speaker prefers to do what it promises, given how its partner will behave if it believes the promise. In a division of labor equilibrium, that means a B-advantaged type prefers to contribute to method B when it expects its partner to give to method A. Similarly, a statement is self-signaling if a speaker wants it to be believed if and only if it is true. In the current example, the B-advantaged type sends a self-signaling message if it would rather its partner believe it is B-advantaged (and thus contribute to method A) than believe it is A-advantaged (and thus contribute to method B). The following proposition summarizes the conditions under which both the self-commitment and self-signaling requirements are satisfied, and therefore a division of labor equilibrium exists. In particular, there is a division of labor equilibrium as long as the marginal effect of each type of contribution on the value of the public good (or the chance of success) is sufficiently large.

Proposition 6.3. *In the two-method game, there exists a division of labor equilibrium if the fol-*

[20] Any influential equilibrium must entail country 2 choosing $x_2 = (1, 0)$ in response to one message and $x_2 = (0, 1)$ in response to the other.

lowing inequalities are satisfied:

$$p(1,1) - p(0,1) \geq \max\left\{\alpha_1, \alpha_2, \beta_1^L\right\}, \tag{6.2}$$

$$p(1,1) - p(1,0) \geq \max\left\{\alpha_1, \beta_1^L, \beta_2\right\}. \tag{6.3}$$

The incentive to misrepresent that impeded communication in prior forms of the model does not show up here, as the informed country's comparative advantage dictates what it wants from its partner. When the conditions of the proposition hold, the A-advantaged type wants its partner to contribute to B, and the B-advantaged type wants its partner to contribute to A. Therefore, neither type has an incentive to mimic the other. The result suggests that uncertainty about comparative advantages across contribution methods should not be a major hindrance to international public goods provision, even when there is no institution in place to manage information. As long as each avenue of contribution is reasonably effective, it is in every state's interest to be honest with its partners about what it is best at. Ordinary diplomacy is thus an effective means of coordinating the division of labor across multiple avenues of contribution to a joint project.

7 Conclusion

This paper has analyzed the role of communication in international public goods problems. The results are largely, though not uniformly, negative. Under a wide set of circumstances, diplomatic talk cannot affect contribution outcomes at all, let alone lead to substantial efficiency improvements. Ordinary diplomacy can be effective only when there is a relationship between a country's private information and what it wants its partners to do. This condition cannot be satisfied in the typical public goods setting, in which all participants strictly prefer that their partners give as much as possible, so diplomatic talk has no effect. These findings buttress institutionalist claims that formal organizations—or other means of un-cheapening diplomatic talk—are necessary to reduce uncertainty in international public goods dilemmas.

The results also call into question the claim that diplomacy can be effective as long as states have shared interests, as previous studies of diplomatic credibility have claimed.

In light of this paper's findings, how can states with common goals best achieve them despite pervasive uncertainty? I have already suggested that that institutional regimes should punish dishonesty rather than failure to contribute, as the most efficient solution to a public goods problem may entail some free-riding. But punishment strategies are not the only way to give states incentives to cooperate in an environment of uncertainty—international institutions should also look to the literature on mechanism design for public goods problems. Institutions with sufficient budgetary resources may be able to coordinate the types of efficient transfer schemes identified by mechanism design theorists (d'Aspremont and Gérard-Varet 1979). Short of that, international organizations may be able to increase public goods provision by establishing routines for sequential giving and by publicizing early contributions (Bag and Roy 2011; Barbieri 2012). Institutions can even improve global welfare simply by drawing more states into the pool of contributors when new problems arise (Bliss and Nalebuff 1984). These strategies for managing asymmetric information in joint projects may be useful as complements or substitutes to the kind of "tit-for-tat" punishments usually discussed in the international institutions literature. A direction for future research would be to design revelation schemes when states have asymmetric intelligence information, which has received much less attention in the mechanism design literature than uncertainty about costs of contribution.

This paper's findings also highlight the need for a more robust body of theory on the role of diplomatic talk in international politics. Besides the notion that communication is effective when states have common interests—a claim that this paper's findings contradict—we have little sense of what kinds of information are exchanged diplomatically or what types of problems diplomacy is best at solving. One particularly important question that remains open is what role communication plays in the formation of coalitions among states. Shifts in international allegiances over time are a major topic of diplomatic history, but existing theories of diplomatic talk only consider relations between pairs of states who know the broad contours of each other’s preferences. Future work on diplomatic communication should consider how states discover their friends and enemies in a world of many actors.

References

Agastya, Murali, Flavio Menezes and Kunal Sengupta. 2007. "Cheap Talk, Efficiency and Egalitarian Cost Sharing in Joint Projects." *Games and Economic Behavior* 60(1):1–19.

Aliprantis, Charalambos D and Kim C Border. 2006. *Infinite Dimensional Analysis: a Hitchhiker's Guide*. Berlin; London: Springer.

Anderson, M S. 1993. *The Rise of Modern Diplomacy, 1450–1919*. London: Longman.

Athey, Susan. 2001. "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information." *Econometrica* 69(4):861–889.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. Basic Books.

Bag, Parimal Kanti and Santanu Roy. 2011. "On Sequential and Simultaneous Contributions Under Incomplete Information." *International Journal of Game Theory* 40(1):119–145.

Baliga, Sandeep and Stephen Morris. 2002. "Co-ordination, Spillovers, and Cheap Talk." *Journal of Economic Theory* 105(2):450–468.

Barbieri, Stefano. 2012. "Communication and Early Contributions." *Journal of Public Economic Theory* 14(3):391–421.

Barrass, Gordon S. 2009. *The Great Cold War: A Journey through the Hall of Mirrors*. Stanford, Calif.: Stanford University Press.

Bergstrom, Theodore, Lawrence Blume and Hal Varian. 1986. "On the Private Provision of Public Goods." *Journal of Public Economics* 29(1):25–49.

Bliss, Christopher and Barry Nalebuff. 1984. "Dragon-Slaying and Ballroom Dancing: The Private Supply of a Public Good." 25:1–12.

Bumiller, Elisabeth and David D Kirkpatrick. 2011. "Libya No-Fly Zone to be Led by NATO." *New York Times* p. A9.

Clarke, Edward H. 1971. "Multipart pricing of public goods." *Public Choice* 11(1):17–33.

Crawford, Vincent P and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6):1431–1451.

d'Aspremont, Claude and Louis-André Gérard-Varet. 1979. "Incentives and Incomplete Information." *Journal of Public Economics* 11(1):25–45.

Erlanger, Steven. 2011. "Confusion Over Who Leads Libya Strikes, and for How Long." *New York Times* .

Farrell, Joseph and Matthew Rabin. 1996. "Cheap Talk." *The Journal of Economic Perspectives* 10(3):103–118.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Fey, Mark, Jaehoon Kim and Lawrence S Rothenberg. 2007. "Pre-Play Communication in Games of Two-Sided Incomplete Information.".

Fudenberg, Drew and Jean Tirole. 1991. "Perfect Bayesian Equilibrium and Sequential Equilibrium." *Journal of Economic Theory* 53(2):236–260.

Green, Jerry and Jean-Jacques Laffont. 1977. "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods." *Econometrica* 45(2):427.

Hildreth, Clifford. 1974. "Expected Utility of Uncertain Ventures." *Journal of the American Statistical Association* 69(345):9–17.

Jervis, Robert. 1976. *Perception and Misperception in International Politics*. 1 edition ed. Princeton, N.J: Princeton University Press.

Keohane, Robert O. 1982. "The Demand for International Regimes." *International Organization* 36(2):325–355.

Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton: Princeton University Press.

Kurizaki, Shuhei. 2007. "Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy." *American Political Science Review* 101(03):543–558.

Kydd, Andrew. 2003. "Which Side Are You On? Bias, Credibility, and Mediation." *American Journal of Political Science* 47(4):597–611.

Martin, Lisa L. 1992. "Interests, Power, and Multilateralism." *International Organization* 46(4):765–792.

McBride, Michael. 2006. "Discrete Public Goods Under Threshold Uncertainty." *Journal of Public Economics* 90(6-7):1181–1199.

Menezes, Flavio M, Paulo K Monteiro and Akram Temimi. 2001. "Private provision of discrete public goods with incomplete information." *Journal of Mathematical Economics* 35(4):493–514.

Michaels, Jeffrey H. 2013. Able but Not Willing: A Critical Assessment of NATO's Libya Intervention. In *The NATO Intervention in Libya Lessons Learned from the Campaign*. pp. 17–40.

Mitchell, Bryan. 2007. "England-based intel center fully operational." *Stars and Stripes* .

Nitzan, Shmuel and Richard E Romano. 1990. "Private Provision of a Discrete Public Good with Uncertain Cost." *Journal of Public Economics* 42(3):357–370.

Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge: Harvard University Press.

Olson, Mancur and Richard Zeckhauser. 1966. "An Economic Theory of Alliances." *The Review of Economics and Statistics* 48(3):266–279.

Ramsay, Kristopher W. 2011. "Cheap Talk Diplomacy, Voluntary Negotiations, and Variable Bargaining Power." *International Studies Quarterly* 55(4):1003–1023.

Samuelson, Paul A. 1954. "The Pure Theory of Public Expenditure." *The Review of Economics and Statistics* 36(4):387.

Sanger, David E. 2004. "Powell Says C.I.A. Was Misled About Weapons." *New York Times* .

Sartori, Anne E. 2002. "The Might of the Pen: A Reputational Theory of Communication in International Disputes." *International Organization* 56(1):121–149.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Reprint edition ed. Cambridge: Harvard University Press.

Seidmann, Daniel J. 1990. "Effective Cheap Talk with Conflicting Interests." *Journal of Economic Theory* 50(2):445–458.

Suleiman, Ramzi. 1997. "Provision of Step-Level Public Goods Under Uncertainty a Theoretical Analysis." *Rationality and Society* 9(2):163–187.

Trager, Robert F. 2010. "Diplomatic Calculus in Anarchy: How Communication Matters." *American Political Science Review* 104(02):347–368.

Trager, Robert F. 2011. "Multidimensional Diplomacy." *International Organization* 65(03):469–506.

A Appendix

A.1 Formal Definition of Equilibrium

An assessment σ is an *equilibrium* if it satisfies the following sequential rationality and consistency conditions for each country i:

- Each type's choice of message is optimal, given the contribution strategies:

$$\mu_i(c_i) \in \arg\max_{m_i \in M_i} \int_{T_j} u_i\left(s_i(c_i, m_i, \mu_j(c_j)), s_j(c_j, \mu_j(c_j), m_i) \,|\, c_i\right) dF_j. \tag{A.1}$$

- In every contribution subgame $\Gamma(m_1, m_2)$, each type's choice of contribution is optimal, given its own beliefs and its partner's strategy:

$$s_i(c_i, m_i, m_j) \in \arg\max_{x_i \in X_i} \int_{T_j} u_i(x_i, s_j(c_j, m_j, m_i) \,|\, c_i) \, d\lambda_i(m_j). \tag{A.2}$$

- The beliefs $\lambda_i(m_j)$ are updated in accordance with Bayes' rule whenever possible (i.e., for all $m_j \in \mu_j(T_j)$).

A.2 Existence of a Babbling Equilibrium

This section proves that a babbling equilibrium exists. The first step is to prove that an equilibrium exists in every subgame in which each country's beliefs about the other's type have interval support.

Lemma A.1. *Consider any contribution subgame $\Gamma(m_1, m_2)$ of the game with incomplete information about costs. If each $\lambda_i(m_j)$ is either a degenerate distribution or a continuous distribution on an interval, then there exists a Bayesian Nash equilibrium of the subgame.*

Proof. First, suppose country 2's belief about country 1's type is a point mass on c_1'; i.e., $\lambda_2(m_1) = \delta_{c_1'}$. (The proof for the opposite case is identical.) Because country 2 believes with certainty that $c_1 = c_1'$, its strategy $s_2(\cdot, m_2, m_1)$ is a best response if and only if

$$s_2(c_2, m_2, m_1) = \max\left\{0, x^*(c_2) - s_1(c_1', m_1, m_2)\right\}.$$

To show that an equilibrium exists, we must find a value of $s_1(c_1', m_1, m_2)$ that is a best response for type c_1' of country 1 when country 2's strategy is defined as above. Let $\Psi(x_1, c_1)$ denote the best response for a given type of country 1 when it expects each type of country 2 to employ a best response to x_1:

$$\Psi(x_1, c_1) = \operatorname*{arg\,max}_{x_1' \in X_1} \left\{ \int_{T_2} p\left(x_1' + \max\{0, x^*(c_2) - x_1\}\right) d\lambda_1(m_2) - c_1 x_1' \right\}.$$

Because contributions above a type's stand-alone contribution are strictly dominated, we have $\Psi(x_1, c_1') \in [0, x^*(c_1')]$ for all x_1. Moreover, by the Berge Maximum Theorem (Aliprantis and Border 2006, Theorem 17.31), Ψ is continuous in x_1. Then, by the Brouwer fixed-point theorem, there exists $\hat{x} \in [0, x^*(c_1')]$ such that $\Psi(\hat{x}, c_1') = \hat{x}$. The following strategies therefore constitute an equilibrium of $\Gamma(m_1, m_2)$:

$$\begin{aligned} s_1(c_1, m_1, m_2) &= \Psi(\hat{x}, c_1), \\ s_2(c_2, m_2, m_1) &= \max\{0, x^*(c_2) - \hat{x}\}. \end{aligned}$$

On the other hand, suppose neither message is fully revealing, so $\lambda_2(m_1)$ and $\lambda_1(m_2)$ are both continuous distributions over an interval. Each country's expected utility in the subgame, holding fixed the strategy of the other country, is

$$Eu_i(x_i \mid c_i) = \int_{T_j} p(x_i + s_j(c_j, m_j, m_i))\, d\lambda_i(m_j) - c_i x_i.$$

Differentiating gives

$$\frac{\partial^2 Eu_i(x_i \mid c_i)}{\partial x_i \partial c_i} = -1,$$

so each country's best response is non-increasing in its type. In addition, because contributions above a type's stand-alone contribution are strictly dominated and x^* is decreasing, we may restrict each country's action space in the contribution stage to $[0, x^*(\underline{c}_i)]$ without loss of generality. Therefore, by Corollary 2.1 of Athey (2001), a pure-strategy equilibrium of the subgame exists. □

This existence result allows us to construct a babbling equilibrium in the usual way, as stated in the following proposition.

Proposition A.2. *A babbling equilibrium exists.*

Proof. Let each μ_i be a babbling messaging strategy, and let every $\lambda_j(m_i)$ be the same as the prior distribution of c_i. Consider any contribution subgame $\Gamma(m_1, m_2)$. By Lemma A.1, there exists an equilibrium of this subgame. Let each country's strategy in every subgame be the same as in the chosen equilibrium. As both countries' beliefs are the same in every subgame, these strategies satisfy sequential rationality. Since the outcome of every subgame is the same, no type of either country has an incentive to deviate from its prescribed message. Finally, beliefs are updated in accordance with Bayes' rule wherever possible. Therefore, the assessment is an equilibrium. □

A.3 Proofs of Stated Results

This section contains the proofs of results stated in the body of the paper.

Proposition 3.1. *Suppose the marginal contribution costs c_1 and c_2 are common knowledge. Without loss of generality, let $c_1 \leq c_2$. If $c_1 < c_2$, the unique Nash equilibrium of the contribution game is $(x_1, x_2) = (x^*(c_1), 0)$. If $c_1 = c_2 = c$, then (x_1, x_2) is a Nash equilibrium of the contribution game if and only if $x_1 + x_2 = x^*(c)$.*

Proof. I begin by ruling out mixed-strategy equilibria. Fix country j's strategy as a mixed strategy given by the probability measure ξ. As p is strictly concave, country i's objective function,

$$Eu_i(x_i) = \int_{X_j} p(x_i + x_j)\, d\xi - c_i x_i,$$

is strictly concave as well (Hildreth 1974). This implies any country has a unique best response to any mixed strategy by its partner, so there can be no mixed-strategy equilibrium. A further consequence of strict concavity of the objective function is that the pure strategy profile (x_1, x_2) is an equilibrium if and only if the first-order optimality conditions

$$x_i \left(p'(x_i + x_j) - c_i \right) = 0,$$
$$p'(x_i + x_j) - c_i \leq 0,$$
$$x_i \geq 0,$$

are satisfied for $i = 1, 2$. The proposition follows immediately. $\square$

Proposition 3.2. *There is no equilibrium that satisfies all of the following criteria:*

1. *There is a pair of messages $m_i, m_i' \in M_i$, with $m_i \neq m_i'$, such that $m_i = \mu_i(c_i)$ for some $c_i \in T_i$ and $m_i' = \mu_i(c_i')$ for some $c_i' \in T_i$.*
2. *For almost all $c_j \in T_j$, $s_j(c_j, \mu_j(c_j), m_i') \geq s_j(c_j, \mu_j(c_j), m_i)$.*
3. *There is a set of types $\tilde{T}_j \subseteq T_j$ such that $\Pr(c_j \in \tilde{T}_j) > 0$ and $s_j(c_j, \mu_j(c_j), m_i') > s_j(c_j, \mu_j(c_j), m_i)$ for all $c_j \in \tilde{T}_j$.*

Proof. Take any assessment that meets the conditions of the proposition. Consider a deviation whereby type c_i of country i sends the message m_i' and then, after receiving country j's message, makes the same contribution as if it had sent the prescribed message m_i. For ease of exposition, let $s_{i|c_i}(c_j) = s_i(c_i, \mu_i(c_i), \mu_j(c_j))$ denote the contribution of type c_i that is realized when its partner sends the message prescribed for type c_j, and define $s_{j|c_j}$ analogously. By construction, $s_{j|c_j}(c_i') \geq s_{j|c_j}(c_i)$ for almost all $c_j \in T_j$, and strictly so on $\tilde{T}_j$. As p is strictly increasing, the difference in expected utility between the deviation and type c_i's proposed strategy is

$$\begin{aligned}
&\int_{T_j} \left[p\left(s_{i\,|\,c_i}(c_j) + s_{j\,|\,c_j}(c_i')\right) - p\left(s_{i\,|\,c_i}(c_j) + s_{j\,|\,c_j}(c_i)\right)\right] dF_j \\
&\quad \geq \int_{\tilde{T}_j} \left[p\left(s_{i\,|\,c_i}(c_j) + s_{j\,|\,c_j}(c_i')\right) - p\left(s_{i\,|\,c_i}(c_j) + s_{j\,|\,c_j}(c_i)\right)\right] dF_j \\
&\quad > 0.
\end{aligned}$$

Therefore, the assessment is not an equilibrium. □

Proposition 4.1. *There is no influential equilibrium in the game with one-sided incomplete information about costs.*

Proof. In an influential assessment, there exist types c_1, c_1' such that $\mu_1(c_1) \neq \mu_1(c_1')$ and

$$s_2(c_2, \mu_2(c_2), \mu_1(c_1')) > s_2(c_2, \mu_2(c_2), \mu_1(c_1)).$$

Since country 2's type is c_2 with probability 1, by Proposition 3.2, such an assessment cannot be an equilibrium. □

Proposition 4.2. *In any equilibrium of the game with one-sided incomplete information about costs,* ex post *total contributions and social welfare are always weakly less than if c_1 were common knowledge. They are strictly less with positive probability if $\underline{c}_1 < c_2 < \bar{c}_1$ and $x^*(c_2) > 0$.*

Proof. Let σ be an equilibrium of the game with one-sided incomplete information about costs. I begin by proving that *ex post* total contributions are weakly less in equilibrium than if c_1 were common knowledge; i.e., that

$$s_1(c_1, \mu_1(c_1), c_2) + s_2(c_2, c_2, \mu_1(c_1)) \leq \max\{x^*(c_1), x^*(c_2)\} \tag{A.3}$$

for all c_1. By Proposition 4.1, in equilibrium country 2 contributes the same amount $\tilde{x}_2 \geq 0$

in response to all messages on the path of play: $s_2(c_2, c_2, \mu_1(c_1)) = \tilde{x}_2$ for all $c_1 \in T_1$. The first-order conditions for an optimal choice by country 1 then give

$$s_1(c_1, \mu_1(c_1), c_2) = \max\{0, x^*(c_1) - \tilde{x}_2\} \tag{A.4}$$

Since any contribution above a country's stand-alone contribution is strictly dominated, $x_2 \leq x^*(c_2)$. The inequality (A.3) follows immediately.

Next, I prove that *ex post* contributions are strictly less for all $c_1 \in (c_2, \bar{c}_1]$ if $\underline{c}_1 < c_2 < \bar{c}_1$ and $x^*(c_2) > 0$. To this end, I will show that these conditions imply $\tilde{x}_2 < x^*(c_2)$. As $\tilde{x}_2 \leq x^*(c_2)$, equation (A.4) gives $s_1(c_1, \mu_1(c_1), c_2) > 0$ for all $c_1 < c_2$. Therefore, the first-order condition for a best response by country 2 cannot be satisfied at $x_2 = x^*(c_2)$:

$$\begin{aligned} &\int_{T_1} p'(s_1(c_1, \mu_1(c_1), c_2) + x^*(c_2))\, dF_1 - c_2 \\ &\qquad < \int_{T_1} p'(x^*(c_2))\, dF_1 - c_2 = 0. \end{aligned}$$

On the path of play, when $c_1 > c_2$, the total contribution is $\tilde{x}_2$ (by equation (A.4)), whereas it would be the strictly greater value $x^*(c_2)$ under complete information.

The final task is to prove that the lower total contribution under incomplete information implies lower *ex post* social welfare. Let c_1 be fixed and let ℓ denote the country with the lower marginal cost of effort,[21] so $c_\ell = \min\{c_1, c_2\}$. Define the *ex post* social welfare function as the sum of the countries' utilities,

$$U(x_1, x_2) = 2p(x_1 + x_2) - c_1 x_1 - c_2 x_2.$$

This is a strictly concave function that is maximized by a contribution scheme in which ℓ spends $x^*(c_\ell/2)$ and the other country spends nothing. Let $(x_1', x_2') = (s_1(c_1, \mu_1(c_1), c_2), s_2(c_2, c_2, \mu_1(c_1))$ denote the contribution scheme realized on the equilibrium path under incomplete informa-

[21] If $c_1 = c_2$, ℓ may be either country.

tion, and let (x_1'', x_2'') be any equilibrium contribution scheme if c_1 were common knowledge. By Proposition 3.1, social welfare under complete information is

$$U(x_1'', x_2'') = 2p(x^*(c_\ell)) - c_\ell x^*(c_\ell).$$

By equation (A.3), $x_1' + x_2' \leq x^*(c_\ell)$, strictly so if $x^*(c_2) > 0$ and $\underline{c}_1 < c_2 < c_1 \leq \bar{c}_1$. This inequality, combined with strict concavity of the social welfare function, gives

$$\begin{aligned} U(x_1', x_2') &= 2p(x_1' + x_2') - c_1 x_1' - c_2 x_2' \\ &\leq 2p(x_1' + x_2') - c_\ell(x_1' + x_2') \\ &\leq 2p(x^*(c_\ell)) - c_\ell x^*(c_\ell) \\ &= U(x_1'', x_2''), \end{aligned}$$

with the final inequality holding strictly if $x_1' + x_2' < x^*(c_\ell)$. *Ex post* social welfare therefore is weakly less under incomplete information, strictly so if the additional conditions of the proposition are met. $\square$

Proposition 5.1. *There is no fully separating influential equilibrium in the game with two-sided incomplete information about costs.*

Proof. Consider a fully separating equilibrium, in which $\mu_i(c_i) = c_i$ for all c_i. By Proposition 3.1, each type's contribution on the path of play is given by

$$s_i(c_i, c_i, c_j) = \begin{cases} x^*(c_i) & c_i < c_j, \\ x^*(c_i) - s_j(c_j, c_j, c_i) & c_i = c_j, \\ 0 & c_i > c_j. \end{cases}$$

This expression is weakly decreasing in c_j. Therefore, if the assessment is influential, it meets the conditions of Proposition 3.2, contradicting the assumption of equilibrium. $\square$

The proof of Proposition 5.2 requires an additional lemma:

Lemma A.3. *If p is quadratic on $[0, \frac{1}{\underline{c}_1} + \frac{1}{\underline{c}_2}]$, then in any equilibrium*

$$s_i(c_i, m_i, m_j) = \max\left\{0, x^*(c_i) - E_{\lambda_i(m_j)}[s_j(c_j, m_j, m_i)]\right\} \tag{A.5}$$

for all $c_i \in T_i$, $m_i \in M_i$, $m_j \in M_j$.

Proof. The quadraticity assumption implies p' is linear on $[0, \frac{1}{\underline{c}_1} + \frac{1}{\underline{c}_2}]$. Moreover, in equilibrium, $s_j(c_j, m_j, m_i) \leq x^*(c_j) < \frac{1}{\underline{c}_j}$ for all $c_j \in T_j$. The first-order condition for $x_i \in (0, x^*(c_i)]$ to be a best response for type c_i in the subgame $\Gamma(m_i, m_j)$ is thus

$$\begin{aligned} 0 &= E_{\lambda_i(m_j)}\left[p'\left(x_i + s_j(c_j, m_j, m_i)\right) - c_i\right] \\ &= p'\left(x_i + E_{\lambda_i(m_j)}[s_j(c_j, m_j, m_i)]\right) - c_i. \end{aligned}$$

This implies $x_i + E_{\lambda_i(m_j)}[s_j(c_j, m_j, m_i)] = x^*(c_i)$, as claimed. □

Proposition 5.2. *If p is quadratic on $[0, \frac{1}{\underline{c}_1} + \frac{1}{\underline{c}_2}]$, there is no influential equilibrium in interval messaging strategies in the game with two-sided incomplete information about costs.*

Proof. Suppose, for the purposes of a proof by contradiction, that there exists an influential equilibrium in interval messaging strategies. Without loss of generality, let country 1 be the influential one, so there exist messages m_1 and m_1' in the range of μ_1 such that $m_1' > m_1$ and $s_2(c_2, \mu_2(c_2), m_1) \neq s_2(c_2, \mu_2(c_2), m_1')$ on a positive-measure subset of T_2.[22] By Proposition 3.2 and the assumption of equilibrium, there is a type $\tilde{c}_2 \in T_2$ that contributes more after receiving the message m_1 than after receiving m_1'; i.e., $s_2(\tilde{c}_2, m_2, m_1) > s_2(\tilde{c}_2, m_2, m_1')$, where $m_2 = \mu_2(\tilde{c}_2)$.

I claim that the expected total contribution in the subgame $\Gamma(m_1, m_2)$ is no greater than that

[22] For expository convenience, I assume throughout the proof that the messaging strategies are defined such that $m_i \in \mathscr{C}_i(m_i)$ for all $m_i \in \mu_i(T_i)$. For any interval partition of T_i, there exists a corresponding messaging strategy μ_i that meets this condition (Fey, Kim and Rothenberg 2007), so there is no loss of generality. Under this condition, each μ_i is weakly increasing, so higher messages correspond to higher-cost types.

in $\Gamma(m_1', m_2)$. Let $\bar{x}_j(m_i, m_j)$ denote country i's expectation of what country j will contribute, given the outcome of the messaging stage:

$$\bar{x}_j(m_i, m_j) = E_{\lambda_i(m_j)}\left[s_j(c_j, m_j, m_i)\right].$$

By Lemma A.3, $s_2(\tilde{c}_2, m_2, m_1) > s_2(\tilde{c}_2, m_2, m_1')$ implies $\bar{x}_1(m_1, m_2) < \bar{x}_1(m_1', m_2)$, as country 2's optimal contribution is inversely related to its expectation of what country 1 will contribute. Again applying Lemma A.3, this in turn gives

$$\begin{aligned}
&\left[\bar{x}_1(m_1', m_2) + \bar{x}_2(m_1', m_2)\right] - \left[\bar{x}_1(m_1, m_2) + \bar{x}_2(m_1, m_2)\right] \\
&\quad = \bar{x}_1(m_1', m_2) - \bar{x}_1(m_1, m_2) \\
&\qquad + \int_{T_2} \max\left\{0, x^*(c_2) - \bar{x}_1(m_1', m_2)\right\} d\lambda_1(m_2) \\
&\qquad - \int_{T_2} \max\left\{0, x^*(c_2) - \bar{x}_1(m_1, m_2)\right\} d\lambda_1(m_2) \\
&\quad \geq \bar{x}_1(m_1', m_2) - \bar{x}_1(m_1, m_2) + \int_{T_2} \left[\bar{x}_1(m_1, m_2) - \bar{x}_1(m_1', m_2)\right] d\lambda_1(m_2) \\
&\quad = 0.
\end{aligned}$$

Therefore, we have

$$\bar{x}_1(m_1', m_2) + \bar{x}_2(m_1', m_2) \geq \bar{x}_1(m_1, m_2) + \bar{x}_2(m_1, m_2). \tag{A.6}$$

To conclude the proof, I show that equation (A.6) contradicts the assumption that $m_1' > m_1$. Take any $\tilde{c}_1 \in \mathscr{C}_1(m_1)$ such that $s_1(\tilde{c}_1, m_1, m_2) = \bar{x}_1(m_1, m_2)$. If $\mathscr{C}_1(m_1)$ is a singleton, then $\tilde{c}_1 = m_1$ by definition; otherwise, $\mathscr{C}_1(m_1)$ is an interval, and the Second Mean Value Theorem for Integrals guarantees the existence of an appropriate value of $\tilde{c}_1 \in \mathscr{C}_1(m_1)$. Let $\tilde{c}_1' \in \mathscr{C}_1(m_1')$

be chosen analogously. Lemma A.3 gives

$$\bar{x}_1(m_1, m_2) = s_1(\tilde{c}_1, m_1, m_2) \geq x^*(\tilde{c}_1) - \bar{x}_2(m_1, m_2), \tag{A.7}$$

$$\bar{x}_1(m_1', m_2) = s_1(\tilde{c}_1', m_1', m_2) = x^*(\tilde{c}_1') - \bar{x}_2(m_1', m_2). \tag{A.8}$$

By definition of a monotone messaging strategy, $\tilde{c}_1' > \tilde{c}_1$. Moreover, by equation (A.8), $x^*(\tilde{c}_1) \geq \bar{x}_1(m_1', m_2) > 0$. As the stand-alone contribution function x^* is weakly decreasing everywhere and strictly decreasing wherever it is positive, $x^*(\tilde{c}_1') < x^*(\tilde{c}_1)$. However, combining equations (A.6)–(A.8) gives

$$x^*(\tilde{c}_1') = \bar{x}_1(m_1', m_2) + \bar{x}_2(m_1', m_2) \geq \bar{x}_1(m_1, m_2) + \bar{x}_2(m_1, m_2) \geq x^*(\tilde{c}_1),$$

a contradiction. □

Strategies and beliefs in the intelligence-sharing game are defined as follows. Country 1's messaging strategy is a function $\mu_1 : Z \to M_1$, where the message space is defined as $M_1 = Z$. Country 1's contribution strategy is a function $s_1 : Z \times M_1 \to X_1$, and country 2's is a function $s_2 : M_1 \to X_2$, where once again $X_1 = X_2 = \Re_+$. Finally, country 2's updated beliefs are λ_2, where each $\lambda_2(m_1)$ is a probability measure on Y.

Proposition 6.1. *If the signal is perfectly informative, then there is a fully separating, influential equilibrium of the intelligence-sharing game.*

Proof. I begin by constructing the claimed equilibrium. Let the messaging strategy μ_1 be fully separating, so $\mu_1(z) = z$ for all $z \in Z$. Let country 2's strategy be

$$s_2(m_1) = \begin{cases} m_1 & m_1 < \frac{1}{c_2}, \\ \frac{1}{c_2} & \frac{1}{c_2} \leq m_1 \leq \frac{1}{c_1} + \frac{1}{c_2}, \\ 0 & m_1 > \frac{1}{c_1} + \frac{1}{c_2}. \end{cases}$$

Along the path of play, let country 1's strategy be

$$s_1(z,z) = \begin{cases} 0 & z < \frac{1}{c_2}, \\ z - \frac{1}{c_2} & \frac{1}{c_2} \leq z \leq \frac{1}{c_1} + \frac{1}{c_2}, \\ 0 & z > \frac{1}{c_1} + \frac{1}{c_2}. \end{cases}$$

Off the path of play (i.e., in subgames in which $m_1 \neq z$), let $s_1(z, m_1)$ be any best response for country 1 to $s_2(m_1)$. Country 1's interim utility function,

$$u_1(x_1 \mid m_1, z) = \mathbf{1}\{x_1 + s_2(m_1) \geq z\} - c_1 x_1,$$

is a step function that is maximized at either 0 or $z - s_2(m_1)$, so a best response always exists. Finally, let country 2's updated beliefs be $\lambda_2(m_1) = \delta_{m_1}$ for all $m_1 \in M_1$. I claim that the assessment $(\mu_1, s_1, s_2, \lambda_2)$ is an equilibrium.

I proceed by showing that the proposed contribution strategies are sequentially rational. Consider country 2's strategy. After receiving the message m_1, country 2's belief about the threshold y is a point mass on m_1. Country 2's interim utility function is thus

$$u_2(x_2 \mid m_1) = \mathbf{1}\{s_1(m_1, m_1) + x_2 \geq m_1\} - c_2 x_2,$$

which is maximized at 0 or at $z - s_1(m_1, m_1)$. Specifically, 0 is a maximizer if country 1's contribution is sufficient for success ($s_1(m_1, m_1) \geq m_1$) or the remaining amount necessary to succeed exceeds country 2's willingness to contribute ($z - s_1(m_1, m_1) \geq \frac{1}{c_2}$). Conversely, if $0 \leq z - s_1(m_1, m_1) \leq \frac{1}{c_2}$, then $z - s_1(m_1, m_1)$ is a maximizer. It is then immediate from the definitions of $s_1(z,z)$ and s_2 that country 2's contribution strategy is optimal, given its beliefs. It is analogous to confirm that country 1's contribution strategy along the path of play ($m_1 = z$) is optimal, given country 2's strategy. In cases off the path of play ($m_1 \neq z$), country 1's contribution is optimal by construction.

The next step is to show that country 1's messaging strategy is incentive compatible—i.e., that country 1 never has an incentive to falsely report the signal it has received. If $z \leq \frac{1}{c_2}$, then country 1's proposed messaging strategy yields its first-best outcome (guaranteed success at no cost to itself), so no profitable deviation is available. If $\frac{1}{c_2} < z \leq \frac{1}{c_1} + \frac{1}{c_2}$, then deviating to any message other than z would cause country 2 to contribute weakly less, so no such deviation may be profitable. Lastly, if $z > \frac{1}{c_1} + \frac{1}{c_2}$, then the most country 1 could induce country 2 to contribute by deviating to a different message is $x_2 = \frac{1}{c_2}$. Even in that case, country 1's best response is to contribute 0, assuring failure of the joint venture and a payoff of 0, the same as under the proposed messaging strategy. Therefore, there is no profitable deviation available.

The final step is to show that country 2's beliefs are consistent with the application of Bayes' rule, which is immediate from the definition of λ_2. □

Proposition 6.2. *If $[0, \frac{1}{c_1} + \frac{1}{c_2}] \subseteq \operatorname{supp} F_y(\cdot \,|\, z)$ for all signals z, then there is no influential equilibrium of the intelligence-sharing game.*

Proof. In an influential equilibrium, there exist signals z' and z'' such that $s_2(\mu_1(z')) < s_2(\mu_1(z''))$. Contributions $x_j > \frac{1}{c_j}$ are strictly dominated for country j, so $s_1(\mu_1(z'), z') + s_2(\mu_1(z'')) \leq \frac{1}{c_1} + \frac{1}{c_2}$. Moreover, $F_y(\cdot \,|\, z)$ is strictly increasing on $[0, \frac{1}{c_1} + \frac{1}{c_2}]$ by construction. Therefore, country 1 can profit by sending the message $\mu(z'')$ when the true signal is z':

$$F_y\left(s_1(z', \mu_1(z')) + s_2(\mu_1(z')) \,|\, z'\right) - c_1 s_1(z', \mu_1(z'))$$
$$< F_y\left(s_1(z', \mu_1(z')) + s_2(\mu_1(z'')) \,|\, z'\right) - c_1 s_1(z', \mu_1(z')).$$

This contradicts the incentive compatibility condition of perfect Bayesian equilibrium. □

Strategies and beliefs in the two-method game are defined as in the intelligence-sharing game. A messaging strategy for country 1 is a function $\mu_1 : T_1 \to M_1$. After receiving country 1's message m_1, country 2 updates its beliefs to $\lambda_1(m_1)$, which is a scalar denoting the probability that $\beta_1 = \beta_1^L$. Country 1's contribution strategy is a function of both its type and the message it sent, $s_1 : T_1 \times M_1 \to X_1$. Country 2's contribution strategy is a function solely of the message it received, $s_2 : M_1 \to X_2$. Before proving the proposition on the existence of a division of labor equilibrium, I state a lemma about when the proposed strategies constitute Nash equilibria of the contribution subgame.

Lemma A.4. *In the contribution stage of the two-method game, suppose β_1 is common knowledge. The strategy profile $\{(x_1^A, x_1^B), (x_2^A, x_2^B)\} = \{(1,0),(0,1)\}$ is a Nash equilibrium if and only if the following conditions are met:*

$$p(1,1) - p(0,1) \geq \alpha_1, \tag{A.9}$$

$$p(1,1) - p(1,0) \geq \beta_2. \tag{A.10}$$

The strategy profile $\{(x_1^A, x_1^B), (x_2^A, x_2^B)\} = \{(0,1),(1,0)\}$ is a Nash equilibrium if and only if:

$$p(1,1) - p(0,1) \geq \alpha_2, \tag{A.11}$$

$$p(1,1) - p(1,0) \geq \beta_1. \tag{A.12}$$

Proof. I will prove the first statement; the proof of the second is analogous. Under the given strategy profile, since $x_2^B = 1$, it cannot be profitable for country 1 to deviate to any strategy with $x_1^B = 1$. Therefore, the only deviation we must consider for country 1 is $(x_1^A, x_1^B) = (0,0)$. This deviation is unprofitable if and only if the condition (A.9) holds. Similarly, because $x_1^A = 1$ under the given strategy profile, the only deviation we must consider for country 2 is $(x_2^A, x_2^B) = (0,0)$. This is unprofitable if and only if the condition (A.10) holds. □

Proposition 6.3. *In the two-method game, there exists a division of labor equilibrium if the fol-*

lowing inequalities are satisfied:

$$p(1,1) - p(0,1) \geq \max\left\{\alpha_1, \alpha_2, \beta_1^L\right\}, \tag{6.2}$$

$$p(1,1) - p(1,0) \geq \max\left\{\alpha_1, \beta_1^L, \beta_2\right\}. \tag{6.3}$$

Proof. I begin by constructing the claimed equilibrium. Let the messaging strategy μ_1 be fully separating, so $\mu_1(\beta_1) = \beta_1$ for all $\beta_1 \in T_1$. Let country 2's strategy be

$$s_2(m_1) = \begin{cases} (1,0) & m_1 = \beta_1^L, \\ (0,1) & m_1 = \beta_1^H. \end{cases}$$

Along the path of play, let country 1's strategy be

$$s_1(\beta_1, \beta_1) = \begin{cases} (0,1) & \beta_1 = \beta_1^L, \\ (1,0) & \beta_1 = \beta_1^H. \end{cases}$$

Off the path of play (i.e., in subgames in which $m_1 \neq \beta_1$), let country 1's strategy be any best response to $s_2(m_1)$; as the game is finite, a best response exists. Finally, let country 2's updated beliefs be $\lambda_2(m_1) = \mathbf{1}\{m_1 = \beta_1^L\}$.

I proceed by showing that the proposed contribution strategies are sequentially rational. The conditions (6.2) and (6.3) imply that equations (A.9)– (A.12) are satisfied. Therefore, by Lemma A.4, country 2's proposed actions are best responses given its beliefs, as are country 1's proposed actions in the subgames on the equilibrium path. In the subgames off the equilibrium path, country 1's actions are best responses by construction.

Next, I show that country 1's messaging strategy is incentive compatible: neither type of country 1 can strictly improve its payoff by deviating from its messaging strategy. Consider the maximal payoff to the B-advantaged type after falsely reporting β_1^H. Since country 2 contributes $x_2^B = 1$ after receiving the A-advantaged message, we only need to consider actions with $x_1^B = 0$. The condition for the deviation to be unprofitable is thus

$$p(1,1) - \beta_1^L \geq \max\left\{p(0,1), p(1,1) - \alpha_1\right\},$$

which follows from equation (6.2) and $\beta_1^L < \alpha_1$. Similarly, the condition for it to be unprofitable for the A-advantaged type to deviate in the messaging stage is

$$p(1,1) - \alpha_1 \geq \max\left\{p(1,0), p(1,1) - \beta_1^H\right\},$$

which follows from equation (6.3) and $\beta_1^H > \alpha_1$.

The final step is to confirm that country 2's updated beliefs are consistent with Bayes' rule wherever possible, which is immediate from the definition of λ_2. □