

The Tragedy of Arming

Working Paper

Andrew J. Coe and Jane Vaynman*

September 20, 2017

Abstract

Arming is puzzling for the same reason war is: it produces outcomes that could instead be realized through negotiation, without the costly diversion of resources arming entails. Despite this, arms control is exceedingly rare historically, so that arming is ubiquitous and its costs to humanity are large. We develop and test a theory that explains why arming is so common and its control so rare. The main impediment to arms control is the need for monitoring that renders a state's arming transparent enough to assure its compliance, but not so much as to threaten its safety. We present evidence that this tradeoff has undermined arms control in three diverse contexts: Iraq's WMD programs after the Gulf War; the superpowers' arms race during the Cold War; and the great power competition in arms in the interwar period. These cases account for 40% of all global arming in the past two centuries.

Word Count: 12,296

*We thank Taylor Dalton for assistance with this research. Comments are welcome and should be sent to andrew.j.coe@gmail.com and jvaynman@gmail.com.

1 Introduction

Why do states spend so much on arms? About 2.4% of the world’s wealth—\$2.4 trillion—was devoted to its militaries in 2012.¹ How much is a mere 2.4 cents out of every dollar? It is enough money to end world poverty, provide bed nets to every person exposed to mosquito-borne disease, identify and treat every person whose infection with HIV would otherwise go undetected or untreated, and increase world spending on research and development by half.²

Despite the immense costs to humanity of arming, it has received very little attention from scholars. Worse, the most fundamental question has been largely ignored: why does arming occur? The answer may seem obvious, but it is not. Arming is mysterious for exactly the reason war is. Why do rational actors employ costly measures to wage their disputes, when negotiation could achieve the same outcome but without the costs?³ In fact arming is even more puzzling than war, since unlike war, arming is not rare. To the contrary, arming is ubiquitous around the world and throughout history.

Most recent theories imply that arming occurs because of asymmetric information: it is assumed that one side cannot perfectly observe the other’s arming decision, and this encourages the latter to arm covertly (Baliga and Sjöström, 2008; Bas and Coe, 2012, 2016*a, b*; Benson and Wen, 2011; Coe, 2017; Debs and Monteiro, 2014; Fearon, 2011, 2016; Jackson and Morelli, 2009; Meierowitz and Sartori, 2008; Powell, 1993).⁴ This certainly accords with

¹This is the most recent figure from the CIA World Factbook.

²In 2012, the World Bank estimated 896 million people lived on less than \$1.90 a day, so that raising all these people to the poverty line would cost no more than \$620 billion per year. Roll Back Malaria estimated in 2015 that 245 million more bed nets were needed for the next several years; according to GiveWell, the total cost of providing each net is \$10. UNAIDS estimated in 2015 that the peak annual cost of a program to identify and treat at least 90% of people with HIV would be \$31 billion. According to UNESCO, worldwide gross expenditures on research and development came to about \$1.5 trillion in 2013.

³See Fearon (1995) for the seminal explication of this inefficiency puzzle of war.

⁴Fearon (2011, 2016); Powell (1993) ostensibly take arming to be perfectly observable, but assume that each side can complete some bounded amount of arming before the other side can observe this and react, which is equivalent to imposing limits on the observability of arming. In earlier theorizing about the security dilemma based on the spiral or deterrence model, arming arises from asymmetric information about preferences rather than actions. However, if arming is perfectly observable, it should still be possible to make an arms control deal, regardless of whether a state is known to be “status quo-oriented” or “revisionist.” See

empirical experience, in which states routinely conceal the details and even overall extent of their arming. However, this doesn't actually answer the question of why arming occurs. Unlike the balance of power or resolve, arming is inherently observable. One side can credibly reveal its arming decision by simply allowing the other side sufficient access to its territory. Thus, most existing models feature arming only because they do not consider the possibility of deals to limit costly arming that utilize monitoring in order to make arming decisions observable.⁵ So, to explain why arming occurs and is so common, we must explain why such deals—which we label “arms control”—are so rare.⁶

We identify a fundamental problem that confronts any attempt to restrain arming, and formalize it in a model of bargaining between two states, where one may arm and the other may undertake a costly response such as counter-arming or war, or the two may instead negotiate an arms control deal. To be viable, any deal must satisfy a transparency requirement and a safety requirement. To ensure the compliance of the side that could arm, the probability that its cheating will be detected must be high enough: monitoring must render this side's arming sufficiently transparent. However, the information revealed by monitoring also cannot give the monitoring side too large a military advantage: that is, the deal must be safe for the arming side.

The problem is that transparency may reveal not only a state's arming decision, but also other information relevant to the balance of power. For example, inspections intended to confirm that a military base does not contain prohibited weapons may also yield information that makes it easier to effectively attack the base or enables the inspecting side to assess other military capabilities. Thus there may be a tradeoff between transparency and safety.

Kydd (1997, 2000) for reviews of this earlier literature and formalizations of these theories.

⁵The only exceptions are Bas and Coe (2016b); Coe (2017), but both models still assume that arming is imperfectly observable, so that a deal to limit arming is often not viable.

⁶The few empirical analyses of arming focus on the effect of domestic politics (Conrad, Kim and Souva, 2013; DiGiuseppe, 2015; Fordham and Walker, 2005; Goldsmith, 2003) or the external security environment (Nordhaus, Oneal and Russett, 2012) on a state's military spending. These studies analyze cross-national differences in these expenditures but do not explain why they are almost always positive and large.

If this tradeoff is severe, it will be impossible to satisfy both requirements at once. Any deal that is transparent enough to assure that one side doesn't arm may shift the balance of power so much that the other side reneges in order to exploit this shift. Any deal that preserves the balance of power well enough to be safe for the arming side may not be transparent enough to assure the monitoring side of its compliance. Thus, no arms control deal may be viable.

Empirically, the transparency-safety tradeoff has severely undermined international attempts at arms control. We show first that the international effort to restrain Iraq from pursuing WMD foundered on this tradeoff. Iraq obstructed the most intrusive international inspections, not because it had something to hide, but because it believed the resulting transparency would be exploited by the US in its ongoing attempts to organize the internal overthrow of Saddam's regime. This in turn meant that the inspections were never sufficient to reassure the world that Iraq had permanently ended its WMD programs, and so contributed to the occurrence of the 2003 Iraq War.

Next, we demonstrate that this tradeoff shaped Cold War negotiations between the superpowers over limiting their arms race. Primary sources indicate that this tradeoff played a role in virtually every negotiation, from the very beginning of the Cold War through to its end. It explains why the earliest attempts to stop the impending arms race failed and why later negotiations over specific limits were so drawn out and detailed. Early agreements were limited to arms that could be monitored unilaterally, precisely because the tradeoff inherent in inspections was severe. Later deals were agreed only once technological innovations made it possible to dampen the tradeoff, by ensuring that inspectors would be able to see only what was necessary to verify the limit, with no other security-relevant information exposed.

Finally, we show that this tradeoff bedeviled inter-war attempts to stop conventional arms competition among the great powers. Great Britain resisted the inspections that would have been required to verify comprehensive limits on arming, because it feared that such inspections would reveal weaknesses in its military capabilities. France insisted on such inspections

because, without them, it could not assure itself of Germany's continued military inferiority. The tradeoff explains why the most successful attempt during this period governed only capital ship construction: these limits could be reliably verified via espionage and so did not require inspections.

The most recent generation of studies of arms control does not consider the transparency-safety tradeoff. Instead it focuses on whether these agreements have escape or withdrawal provisions (Koremenos, 2001), their level of legalization (Kreps, 2016), and the centralization of negotiation (Verdier, 2008), monitoring (Dai, 2002, 2007), and enforcement (Coe and Vaynman, 2015). Vaynman (2017) examines why only some agreements have verification provisions, but relies on a more general conception of the costs of verification. Our transparency-safety tradeoff generates a new "disclosure dilemma" that is analytically similar to those identified by Carnegie and Carson (2017*b,a*), wherein a state's disclosure of evidence that another has cheated on an agreement may undermine either other states' incentives to comply or the first state's future ability to detect cheating. Relatedly, Lindsey (2015) shows that a state may not disclose military strength, even if this leads to war, because revealing it would lead its opponent to take effective counter-measures.

An earlier generation of work on arms control ignored or underplayed the importance and ubiquity of the transparency-safety tradeoff. One branch of this literature investigated the optimal design of verification, but took for granted that monitoring would be imperfect and so ignored our tradeoff (Downs and Rocke, 1990; Kilgour, 1994; Wittman, 1989). A second branch acknowledged that this tradeoff sometimes played a role, but instead viewed domestic politics as the most critical determinant of arms control (Dunn, 1990; Gallagher, 1997, 1999; Goldblat, 2002; Krass, 1985).

2 Model Setup

A (the monitoring state, referred to as feminine) and B (the arming state, masculine) bargain over revisions to a prior division of a composite of disputed issues, represented by $[0, 1]$. The two players have linear preferences over the interval, with A favoring settlements closer to 1 and B favoring those closer to 0, and discount future payoffs by a factor $\delta < 1$.

In the first of infinitely many discrete periods of time, A chooses either to take a costly action against B , which ends the game as described subsequently, or to make a peaceful offer of a settlement for that round. If A makes an offer, B must first choose whether to open or close himself to monitoring. He can then reject A 's offer, ending the game, or accept it, in which case it is implemented for that period.⁷ If B accepts A 's offer, then B chooses whether to invest in arming, which imposes a cost on B of $k > 0$, and succeeds with probability $\lambda > 0$ and otherwise fails. If the investment succeeds, B is newly-armed and this immediately becomes common knowledge.⁸ If it fails, then B must invest again in the future to have a chance of its efforts succeeding. The period then ends. We assume that the length of a period is relatively short, so that k and λ are relatively small: the expected arming that will occur in a single period of investment is modest, though over many periods of repeated investment it may become quite large. This ensures that the discrete-time structure of the game does not artificially prevent A from being able to react quickly to B 's arming. We also assume that B begins the game in the condition of being closed to monitoring.

The structure of the game in subsequent periods depends only on whether B became newly-armed at some point in the past. Once B has done so, subsequent periods entail only repeated bargaining: in each round, A has only to take her costly action or make an offer which B then accepts or rejects. If B has not done so, then each subsequent period until he

⁷The results depend only on whether a surplus exists, not on how it is divided between the players. Thus the choice of a take-it-or-leave-it protocol is an innocuous simplification.

⁸ B has a strong incentive to credibly reveal its new power in order to gain bargaining leverage over A .

does is the same as the first, except that it begins with the receipt of new information by A on whether B invested in the previous round, in the form of a public signal.⁹ If B is closed to monitoring and has invested, then with probability τ_c the signal indicates that he did, and otherwise indicates that he did not. If B is open to monitoring and has invested, then with probability $\tau_o \geq \tau_c$ the signal indicates that he did, and otherwise indicates that he did not. If B did not invest, then the signal indicates that he did not regardless of whether he was open or closed to monitoring. Thus, A 's information on B 's investment is prone to false negatives but not false positives.¹⁰ All choices except for B 's investment, and all parameters of the game, are common knowledge.

If A takes her costly action or B rejects an offer, the game ends in costly conflict. The expected value of this outcome for each state depends on whether B is newly-armed and, if he is not, whether he is open to monitoring. If B is newly-armed, the values are W_n^A and W_n^B respectively. If B is not newly-armed, the values are W_o^A and W_o^B if he is open to monitoring and W_c^A and W_c^B if he is not. We assume $W_o^A \geq W_c^A > W_n^A$ and $W_n^B > W_c^B \geq W_o^B$. It will sometimes be convenient to treat W_o^A and W_o^B as functions of τ_o , with $W_o^A(0) = W_c^A$ and $W_o^B(0) = W_c^B$. We assume that $W_o^A(\cdot)$ is nondecreasing and $W_o^B(\cdot)$ is nonincreasing: B opening up to increasing levels of monitoring does not improve his value from conflict or lessen A 's. All these values are assumed to be non-negative.¹¹ Conflict is costly: the value destroyed by it is $D_x^w \equiv \frac{1}{1-\delta} - W_x^A - W_x^B > 0$, for any $x \in \{n, o, c\}$, with $D_o^w \leq D_c^w$.

A 's costly action is intended to represent any unilateral response to B 's anticipated or actual arming. It might be a decisive invasion or merely limited strikes; covert intervention to overthrow B 's regime or the imposition of sanctions to constrain his military; or counter-

⁹In the equilibria studied here, A never has any incentive to conceal this signal. Empirically, A would have strong incentives to credibly reveal information on B 's investment to build support for action.

¹⁰Allowing for a modest probability of false positives would not change our results qualitatively. The higher this probability becomes, the harder it is to sustain a deal, since compliance might still lead to punishment after a false positive signal.

¹¹This is necessary only to avoid a surfeit of quantities like $\max\{W_o^B, 0\}$ in the exposition.

arming to match B 's arming. All of these actions are costly and intended to preserve A 's power relative to B 's, and so are strategically equivalent from the perspective of our theory. Leaving A 's particular response unspecified enables us to apply the model to empirical cases in which A would launch a war to prevent arming, and also to situations in which A would instead simply arm herself. In keeping with the variety of possible forms game-ending conflict may take, we will refer to it interchangeably as “war” or an “arms race.”

The main novel feature of this model is that the observability of arming is endogenous: the arming side can choose to be closed or open to monitoring. If it is closed, then the other side's ability to observe arming is limited to unilateral intelligence-gathering. If it is open, then this ability is supplemented by bilateral cooperation. This might take the form of the arming state admitting inspectors, allowing overhead surveillance, or voluntarily reporting information. Arming openly obviously means that it is more likely to be detected ($\tau_o \geq \tau_c$). It may also expose information other than merely whether the monitored state is arming, which may disadvantage it and advantage the other state in an arms race or war ($W_o^B \leq W_c^B$ and $W_o^A \geq W_c^A$). It may also reduce the costs of this outcome, as it makes possible a more precise reaction by A to B 's arming ($D_o^w \leq D_c^w$). Inspectors may include spies from the monitoring side who can recruit sources or organize coup attempts, or gather targeting intelligence for limited strikes to stop arming. Overhead surveillance may ascertain current military forces' quality and quantity and discover weaknesses. Voluntary reports may enable unintended inferences about undisclosed information. That the observability of arming may affect the two sides' prospects in conflict is the source of the transparency-safety tradeoff and the key to our results.

Our model takes the level of monitoring to be a unilateral, binary choice by the arming side (either τ_c or τ_o). More realistically, this side could bargain with the other about exactly how open to be—where inspectors can go, with how much notice, and so on—and so choose the degree of openness τ_o . However, allowing this bargaining to occur would not change our

key finding, which is that, under some conditions, a deal would not be viable for *any* degree of openness.

In other respects, our setup is similar to that of models of nuclear proliferation, such as [Bas and Coe \(2016a,b\)](#); [Benson and Wen \(2011\)](#); [Coe \(2017\)](#); [Debs and Monteiro \(2014\)](#). These models are built to analyze *whether* a state arms: for example, whether it pursues nuclear weapons or some other new military technology. For tractability, they assume a binary choice over the level of spending on arming—some or none—but allow the resulting arming to be both delayed and uncertain. By contrast, models of conventional arming are designed to understand *how much* a state arms: e.g., how many soldiers it conscripts or ships it builds. These latter models allow for a continuous choice over the level of spending but assume that it has an immediate, certain effect on the balance of power. Because our focus is on whether two states can avoid arming, we employ a setup of the former type, but we believe qualitatively similar results could be obtained using the other type.

3 Analysis

We first establish the “no-deal” equilibrium, governing behavior in the absence of a deal to stop arming.¹² This forms the backdrop for our analysis of the possibilities for arms control deals and how these are affected by the transparency-safety tradeoff. To render its effects more starkly, we first study the situation when there is no tradeoff, and then consider what happens when it does occur. Proofs of all results appear in the [online appendix](#).

Throughout, we will assume that B finds investing in arming attractive enough to do so in the absence of a deal. If this were not true, then B would never invest even if A offered him the bare minimum settlements, no arming or conflict would occur, and there would be

¹²Equilibrium refers to a Perfect Bayesian Equilibrium. No-deal and deal equilibria are defined formally in the [online appendix](#). Informally, a deal is an equilibrium in which neither investment nor costly conflict occur. We will explain subsequently why such an equilibrium entails a deal being made.

no need for any deal. We discard this uninteresting case.

Assumption 1. *In the absence of a deal, B would invest in arms, given the chance: $k < \delta\lambda [W_n^B - W_c^B]$.*

Assumption 1 means that any no-deal equilibrium will be inefficient: it will feature either costly investment in arming by B , or a costly action by A to preempt this.

Proposition 1. *In the absence of a deal, B remains closed to monitoring. War or an arms race occurs if and only if $W_c^A + \overline{V_c^B} > \frac{1}{1-\delta} - \frac{k}{1-\delta(1-\lambda)}$, where $\overline{V_c^B} \equiv \max \left\{ W_c^B, \frac{-k+\delta\lambda W_n^B}{1-\delta(1-\lambda)} \right\}$.*¹³

To understand this result, observe that in the absence of a deal, if A does not take her costly action, then she will only offer to B the minimum settlement he would accept, and keep the rest for herself. Offering more will bring no benefit to A and leaves her with less of the disputed stake; offering less would be rejected by B . If B were newly-armed, his expected value from conflict would be higher, and A would be forced to offer him a more generous settlement to avoid it than was required before he armed. If B did not invest, then he would never become newly-armed and there would be little for A to fear. However, because B does invest in equilibrium, A has cause to worry that his effort will eventually succeed, after which A will have to concede more of the disputed stake. Moreover, the investments themselves are also costly, lessening the total value to be divided between the two sides.

There are thus two distinct motives for conflict: to lessen an adverse shift in the balance of power due to B 's arming, and to avoid the costs of that arming. These motives derive from a linked pair of commitment problems. B cannot commit not to take advantage of being newly-armed to extract more generous offers from A . Given that better offers would result from being newly-armed, in the no-deal equilibrium B also cannot commit not to seek arms. This leads A to expect that peace will be costly due to B 's investments and unfavorable

¹³Here we ignore the knife-edge case of equality, in which both a conflictual and a peaceful equilibrium exist in the absence of a deal. We also assume that if B is indifferent, he chooses to be closed to monitoring.

due to the potential for a shift in power. If the combination of these two motives is strong enough, war or an arms race will result; otherwise A will simply tolerate B 's arming.

Remarkably, B 's decision to conceal his arming has nothing to do with disguising his arming or intention to do so. Indeed, because in equilibrium it is common knowledge that B will invest, there is never any uncertainty about his doing so, whether he is closed to monitoring or not. Concealment here is instead entirely about preserving advantage in conflict. Because opening to monitoring undermines B 's prospects in conflict ($W_c^B \geq W_o^B$), doing so would only lead to A making less generous offers or resorting to conflict, with B doing worse either way. Thus, B remains closed to avoid A exploiting his openness.

The only way for the two sides to avoid the costs of investment or conflict is to make an arms control deal. In such a deal, B agrees not to arm, and possibly to being open to monitoring, and in exchange A agrees to make more generous offers to B . If either side reneges, and this cheating is detected, then the two sides revert to a punishment equilibrium, such as the no-deal equilibrium (with continuation values denoted V_c^A and V_c^B). First consider the possibility of a deal in which B remains closed to monitoring, termed a “closed deal.” When B is closed, A can only detect B 's investment in arms through means that do not require B 's cooperation, such as recruiting human spies, eavesdropping on electronic communications, and overhead imagery.

Proposition 2. *Let $V_p^A \equiv V_c^A$, $V_p^B \equiv V_c^B$, and $S \equiv \frac{1}{1-\delta} - V_p^A - V_p^B$. There is a deal in which B is closed to monitoring if and only if $\delta[\lambda + \tau_c(1 - \lambda)]S \geq -k + \delta\lambda[W_n^B - V_p^B]$.*

For a deal to be viable, both sides must see compliance with it as offering at least as high a value as reneging on it. Because a deal enables the two sides to avoid the costs of investment or conflict, it creates a surplus that A could use to encourage B 's compliance with the deal, while still leaving both sides better off than each would be in the punishment equilibrium. This surplus must be large enough to overcome B 's temptation to cheat by

covertly investing in arms, in the hopes of escaping detection long enough to become newly-armed. The inequality in the proposition corresponds to this requirement. The right side is B 's temptation to renege by seeking arms: the expected benefit of investment, relative to what B would receive if caught and punished ($\delta\lambda [W_n^B - V_p^B]$) minus its cost (k). The most A can offer to reward B 's compliance is the value left after both sides receive at least what they would in the punishment equilibrium (S). Offering any more would leave A with less value than she would receive if the deal ended, and so would lead to her reneging. A can only threaten to withdraw this reward if B is caught cheating, either by detection prior to becoming newly-armed or by the revelation that he has become newly armed. Thus this punishment is weighted by the chance that B 's investment works (λ) plus the chance that it doesn't and is detected by A ($\tau_c(1 - \lambda)$). If the expected value of this exceeds that of investing, then there is a closed deal that B would not cheat covertly on, and that left both sides better off than they would be in the deal's absence.

When a deal is viable, it enables the two sides to escape the commitment problems that lead to inefficient outcomes in a deal's absence. The reward for compliance and the corresponding punishment for detected cheating renders B able to commit not to invest in arms, so that investment will not occur and preventive attack or counter-arming is unnecessary. However, the viability of a deal is impeded by asymmetric information about B 's decision to invest.

Corollary 1. *A closed deal is possible if and only if unilateral monitoring is good enough ($\tau_c \geq \underline{\tau}$).*

Clearly, the higher the level of monitoring, the easier it is to deter B from cheating because cheating is more likely to incur punishment. The surer punishment is, the smaller the reward A must offer to ensure compliance, and the more likely a deal is to be viable. Intuitively, as long as A would detect B 's investment reliably enough and quickly enough, then any attempt to secretly arm under a deal would be so unlikely to succeed prior to

getting caught and punished that it would not be worth trying. There will therefore be some minimum level of monitoring (τ) that renders B 's choice to invest in arms just transparent enough to assure A of his compliance with a deal. We term the need for monitoring to be at least this good the “transparency requirement.”

Unfortunately, A 's unilateral monitoring of B 's arming (τ_c) might not be good enough to meet the transparency requirement. When this is true, B cannot commit to refrain from investing in arms, so that A will be unwilling to extend any concessions to B , and no deal is possible. However, if B opened to monitoring, A 's unilateral monitoring would be supplemented by B 's cooperation, which might include allowing inspectors onto its territory, tolerating overhead surveillance, or providing relevant information. As a result, A would be better able to detect an investment in arms. Consider next the possibility of such an “open deal.”

Proposition 3. *Let $V_p^A \equiv \max\{W_o^A, V_c^A\}$, and $V_p^B \equiv W_o^B$ if $W_o^A > V_c^A$ and $V_p^B \equiv V_c^B$ otherwise. There is a deal in which B is open to monitoring if and only if $\delta[\lambda + \tau_o(1 - \lambda)]S \geq -k + \delta\lambda[W_n^B - V_p^B]$ and $V_p^A - \frac{k}{\delta} + \lambda W_n^B + (1 - \lambda)V_c^B \leq \frac{1}{1 - \delta}$.*

Just as with a closed deal, both sides must see abiding by the open deal as preferable to any of their options for renegeing on it in order for it to be viable. The first inequality assures that B would not choose to cheat covertly on the deal by investing in arms while remaining open. It is identical to the condition for a closed deal, except that τ_c is replaced by τ_o to reflect that B is open to monitoring here.

Unlike in the closed deal, B here also has the option to renege *overtly* by first closing himself to monitoring and then investing in arms (which would cost k now but yield W_n^B next period if the investment succeeded and V_c^B if it failed). For her part, A could renege by making a lower offer than the deal stipulates, causing both sides to revert to the no-deal equilibrium (yielding V_c^A for A) or by taking her costly action (W_o^A). Thus, a deal must give each side a value sufficient to match that of these options. The second inequality in the

proposition corresponds to these requirements: the values the deal must give to A and B should not sum to more than the total value available ($\frac{1}{1-\delta}$).

These two conditions imply a subtle obstacle to the viability of any open deal, which we call the “transparency-safety tradeoff.” The problem is that the level of openness (τ_o) affects both how easy it is for A to assure herself that B is not covertly cheating on the deal, and also how the two sides will do in costly conflict. To understand how the tradeoff affects the prospects for an open deal, first suppose that B opening to monitoring does not affect either side’s prospects in costly conflict ($W_c^A = W_o^A$ and $W_c^B = W_o^B$). Then there is no transparency-safety tradeoff, and it is relatively easy for the two sides to make a deal.

Corollary 2. *If there is no transparency-safety tradeoff, then there is always a level of open monitoring high enough to support a deal, and this level never need be perfect ($\underline{\tau} \leq \tau_o < 1$).*

Cooperation improves the efficacy of monitoring relative to its unilateral level. By giving A ’s inspectors access to relevant facilities, undertaking certain activities without concealment from A ’s intelligence capabilities, or reporting specific information to A , B can raise the probability that A would detect an investment in arms. With sufficient cooperation—such as allowing inspectors access to any place at any time without notice or delay—this probability would be 1, so that A would be certain to detect cheating. However, because B does strictly better under the deal than he would in its absence, such perfect monitoring is never actually required to support a deal.

This means that the transparency requirement can always be satisfied in an open deal. Even if A ’s unilateral monitoring is not good enough to support a closed deal, B can easily dispel the asymmetric information about his arming that impedes such a deal, by opening sufficiently to monitoring. In the absence of the transparency-safety tradeoff, B is happy to do so because the deal lives him strictly better off. The key here is that increasing transparency by opening to monitoring has no downside for B .

Now suppose that opening to monitoring not only improves A 's ability to detect B cheating but also her prospects in war or mutual arming. Then there is a transparency-safety tradeoff, and a mutually-acceptable deal is less likely to exist.

Corollary 3. *Let $\bar{\tau}_o$ be the largest value of τ_o such that $W_o^A(\tau_o) - \frac{k}{\delta} + \lambda W_n^B + (1-\lambda)V_c^B \leq \frac{1}{1-\delta}$. If the transparency-safety tradeoff is mild (i.e., $\underline{\tau} \leq \bar{\tau}_o$), then a deal will be viable under any level of openness in $[\underline{\tau}, \bar{\tau}_o]$. If the tradeoff is severe (i.e., $\underline{\tau} > \bar{\tau}_o$), then no deal exists under any level of open monitoring, even perfect monitoring.*

With a transparency-safety tradeoff, an open deal must now satisfy two requirements. First, the level of openness must render B 's choice to invest in arms transparent enough to be quickly detected, so that B would be deterred from cheating on the deal. This is the transparency requirement mentioned earlier, and it applies whether there is a transparency-safety tradeoff or not: τ_o must be at least $\underline{\tau}$. Second, the level of openness must not expose so much militarily useful information that A would rather exploit this information by attacking or counter-arming than abide by the deal; that is, the openness has to be safe enough for B . This is the “safety requirement”: τ_o can be no more than $\bar{\tau}_o$.

Obviously, these two requirements are in tension. If the tradeoff between them is mild—safety does not decrease much as transparency increases—then there will be a range of levels of openness ($[\underline{\tau}, \bar{\tau}_o]$) that would all support a mutually-acceptable deal. In this range, monitoring is transparent enough to convince A of B 's compliance, but not so revealing as to be unsafe for B . When the tradeoff is severe, so that safety declines rapidly as transparency increases, there is no level of openness that would both assure A of B 's compliance and assure B of A 's.

The transparency-safety tradeoff makes a deal under open monitoring harder by giving rise to a new commitment problem. Because openness improves A 's value from costly conflict, she may be unable to commit not to exploit it by going to war or counter-arming. This commitment problem leads B to refuse too-open monitoring, which creates asymmetric

information about his arming. The combination of these bargaining problems can make a deal impossible and so lead to costly conflict.

This result builds upon the conclusions of prior theories of arming. In perhaps the earliest bargaining model of arming, [Powell \(1993\)](#) shows that the costs and shifts in power that attend arming explain why arming, or war to prevent it, may occur. [Debs and Monteiro \(2014\)](#) extended this result by showing that arming or war to prevent it could only happen in the presence of asymmetric information about the choice to arm. Without asymmetric information, one side's arming could be perfectly observed by the other side, which could therefore threaten certain punishment, deterring the first side from arming and rendering war unnecessary. Here, we have shown that asymmetric information about arming can only lead to war if the monitoring side is unable to commit not to exploit cooperative monitoring for its own military advantage. In the absence of this commitment problem, there is always some level of monitoring that would support an arms control deal acceptable to both sides, so that any asymmetric information is not enough to cause arming or war.

The results presented here are easily translated into observable implications. Empirically, there are many dimensions along which arming could be controlled. Restrictions could be agreed on the quantity, accuracy, or range of missiles; the number and tonnage of warships; the testing of nuclear warheads; or the size of ground forces. And there are similarly many dimensions to open monitoring: the frequency of inspections; the sites inspectors can access; the number of overflights allowed. The intensity of the transparency-safety tradeoff may vary among these dimensions of arming and monitoring, between countries, and across times.

Our theory implies that, if and only if unilateral monitoring renders a dimension of arming transparent enough to convince the monitoring side that the arming side's compliance can be assured, there will be a closed deal to limit this dimension of arming. If instead the monitoring side believes a certain level of cooperation is required to assure sufficient transparency, and this level is also seen as safe enough by the arming side, there will be

an open deal. Finally, if the cooperation necessary to assure transparency to one side is seen as too threatening to the safety of the other, there will be no deal at all. In sum, the transparency-safety tradeoff should determine whether an agreement is made, what kind of monitoring it features, and what dimension of arming it limits.

4 Empirical Evidence

We proceed to demonstrate that the tradeoff between transparency and safety is a serious and common impediment to making arms control deals empirically, and that policymakers often recognize it as such. Because we cannot feasibly test our theory on all the cases to which it should apply, we chose to examine three that span a variety of contexts—contemporary and historical, nuclear and conventional weapons, major and minor powers, bilateral and multilateral. We first investigate the negotiations between the United States and Iraq over the latter’s pursuit of weapons of mass destruction (WMD), showing that more transparent monitoring severely threatened the safety of Iraq’s regime, preventing any deal from being viable. We then turn to the Cold War, examining the negotiations between the US and USSR over limiting the nuclear arms race between them, where we find that our tradeoff was a constant concern for both sides and affected the kinds of monitoring and limits that could be agreed. Finally, we look back to efforts to constrain the conventional arms competition among the great powers between the World Wars, showing that these were limited to certain naval arms because of the transparency-safety tradeoff. These cases illustrate the ways in which the tradeoff shapes whether an agreement is made, what kind of monitoring it features, and what exactly is limited under it.

However, our chosen cases are not merely illustrative: they represent a large portion of *all* arming. The main parties to the interwar arms talks—the US, UK, France, Germany,

Italy, and Japan—account for over half of global military spending from 1919 to 1938.¹⁴ The US and USSR accounted for almost two-thirds of total military spending from 1946 to 1991. Taken together, these cases make up almost 40% of the world’s military spending for the last two centuries.¹⁵

4.1 No Deal: Iraq and the United States, 1990–2003

The United States imposed sanctions on, repeatedly undertook limited strikes against, and eventually invaded Iraq to overthrow its regime in 2003, all largely for the purpose of ending its pursuit of weapons of mass destruction. We now know that Iraq in fact had no WMD and had suspended its programs by 1996, but that it planned to resume these programs once sanctions had been ended (Iraq Survey Group, 2004, Key Findings, 24, 44, 49, 51). The costs of those programs and of the resulting sanctions, strikes, and war were tremendous for both the US and Iraq. Both sides would have been much better off had they instead struck a deal, in which Iraq disarmed and enabled the US to verify this, and in exchange the US did not sanction, strike, or invade Iraq. Why wasn’t a deal made?

We will argue that no deal was possible because of a severe transparency-safety tradeoff. Iraq’s record of successfully concealing its programs from unilateral monitoring demonstrated that intelligence alone could not assure the US of Iraq’s disarmament. However, inspections intrusive enough to meet this transparency requirement would have exposed Iraq’s regime to very serious risk of being internally overthrown. We draw on the arguments and evidence assembled for this claim in Koblentz (Forthcoming), combining these with our own arguments about unilateral monitoring to explain why no deal was viable.¹⁶

¹⁴All figures are adjusted for inflation and computed from the National Military Capabilities dataset, which covers almost all countries from 1816 to 2012.

¹⁵This excludes spending by the belligerents during the world wars from the total. Presumably an arms control among the belligerents was ruled out by total warfare rather than by our tradeoff.

¹⁶The conventional wisdom for why Iraq did not fully reveal its lack of WMD is that it sought to maintain ambiguity in order to deter Iran. The origin of this theory is a statement made by Saddam Hussein after his capture. Koblentz (Forthcoming) argues that Saddam had incentives to lie about the true cause of his

No deal under purely unilateral monitoring was possible. Iraq had successfully hidden the extent and progress of its nuclear program from the US prior to the Gulf War ([Iraq Intelligence Commission, 2005](#), 53), and for years after continued to conceal elements of this program and most of its biological weapons program ([Iraq Survey Group, 2004](#), 44–51). These were discovered only after the chance defection of Saddam’s son-in-law, Hussein Kamel, as the result of a family dispute ([Iraq Survey Group, 2004](#), 45–47). Thus the US had little cause for confidence that, in the absence of highly transparent cooperative monitoring, it would be able to detect Iraq’s cheating reliably enough to support a deal. Indeed, the US intelligence community’s conclusion in 2002 (and 1999 and 1997) that Iraq was still pursuing WMD was based almost entirely on presumption: given Iraq’s skill at concealment, the lack of direct evidence for continuing programs was not unexpected by analysts ([Iraq Intelligence Commission, 2005](#), 9–10, 46–47, 49, 81–82, 114–115, 154–156, 169–170). The establishment of UNSCOM was motivated by the recognition that Iraq’s WMD disarmament could not be assured without cooperative monitoring.

Thus, only cooperative monitoring could meet the transparency requirement for a viable deal. [Debs and Monteiro \(2014, 25\)](#) argues that inspections were incapable of “proving the negative”—that Iraq was not still pursuing WMD—and other rationalist accounts of the origins of the Iraq War assume that this monitoring was necessarily flawed ([Coe, 2017](#); [Harvey, 2011](#)). However, there was a sufficiently intrusive set of inspections that, if permanently implemented, would have satisfied the US as to Iraq’s abandonment of WMD. In principle, the US could have blanketed Iraq with cameras and inspectors—able to watch every possible facility continuously—and the costs of this would have been small relative to the costs of war. In practice, the Iraq Survey Group was able to do so at modest cost, even given an ongoing insurgency and resistance from the former regime ([Iraq Survey Group, 2004](#), Scope Note).

WMD obfuscation, and shows there is no evidence that corroborates this theory and ample evidence that contradicts it. Koblentz uses this case to demonstrate that concerns about internal rather than external security may dominate the foreign policy choices of autocratic states.

Transparency sufficient to satisfy the US was therefore certainly *possible* with cooperative monitoring.

The problem was that such high transparency would seriously endanger Iraq's regime and the survival of Saddam personally. Over the course of this period, the US covertly helped some of Iraq's officials to organize a coup, and covertly trained and equipped domestic groups opposed to the regime (Koblentz, *Forthcoming*, 23–24).¹⁷ These efforts and those of independent elements within Iraq resulted in a number of popular uprisings, coup plots, and assassination attempts against Saddam (8–9). Though all of these failed to overthrow Iraq's regime, they nonetheless demonstrated a clear and dire threat to it (15–17).

The transparency of Iraq's WMD activities and the safety of its regime were closely connected (3, 22–26). The CIA penetrated the inspection teams and reportedly used inspection visits to coordinate with coup-plotters and opposition elements (24–25). Saddam's presidential palaces were numerous and large enough to hide serious WMD activity, so that transparency required them to be inspected.¹⁸ However, these were also places where Saddam lived, and inspections of them would have revealed valuable information about Saddam's security measures and how to defeat them (24–25). Iraq's Special Security Office was tasked with managing the concealment of Iraq's WMD programs, and thus its facilities and personnel would have to be subject to inspection and interview to confirm that those programs had ended (17, 19, 20). But the very same office was also in charge of Saddam's personal security and transportation (16–17). If these sites had been completely open to inspection, it seems very unlikely the US would not have used those inspections as an opportunity to gather intelligence on the regime's vulnerabilities.

Saddam was keenly aware of the tradeoff between transparency and safety involved in the inspections, and his behavior was driven by it (3–4, 22–26). In captured audiotapes of Sad-

¹⁷We henceforth use only page numbers in citing this article.

¹⁸At least 48 of these palaces were built after the Gulf War, with several occupying multiple square kilometers of grounds (Department of State, 1999).

dam's meetings with senior officials, those present discuss the possibility that acquiescing to the latest request from the inspectors would threaten the regime's internal security (22–23). As a result, the most contentious issues in negotiations between Iraq and the international community over inspections were those for which the tradeoff was most severe, especially inspections of Saddam's palaces and of the SSO (27–28). Almost all of the instances in which an inspection visit was obstructed involved such sites (9–12, 21–22).

The US was largely unaware of how pressing this tradeoff was for Iraq (3–4). High-level officials interpreted Iraq's refusal to accept fully transparent inspections as evidence that it had WMD activity to hide, even as some intelligence analysts pointed out that Iraq's truculence might be additionally motivated by concerns for the regime's safety (27–30). Even after the war, investigations into Iraq's motives for resisting the inspections misattributed this behavior to the need to deter Iran or vaguely-defined domestic enemies (2, 5–12).

Given a severe transparency-safety tradeoff, our theory stipulates that no deal would be mutually acceptable. And indeed, over the course of this period the US constantly demanded more transparency from Iraq, and Iraq resisted mainly in order to protect the regime's safety (4), but no deal was made.

4.2 Difficult Negotiations: Verification in the Cold War

The US and USSR undertook the most expensive arms race in human history, judged by absolute military spending. Moreover, since their arsenals were the most destructive ever created, and some weapons generated incentives to strike first, their arming posed an unprecedented risk to humanity's survival. In their attempts to negotiate constraints on their arming in order to reduce its costs and risks, both sides recognized the transparency-safety tradeoff, and it was crucial to their evaluation of mooted limits and monitoring. The US typically viewed cooperative monitoring as necessary to verify compliance, but the USSR saw the associated inspections as likely to be dangerously revealing, so that the tradeoff

was generally severe and open deals not viable. This pattern is present in the primary sources for every major negotiation, but for brevity we discuss just three here. The Freeze negotiations failed while the subsequent, similar Strategic Arms Limitation Talks (SALT I) succeeded because unilateral monitoring improved enough in the interim to make a closed deal viable. The Intermediate Nuclear Forces (INF) talks resulted in an open deal only once a technological innovation was devised to dampen the transparency-safety tradeoff.

Archival evidence shows that US policymakers evaluated possible agreements in terms of the tradeoff from the very beginning of the Cold War. A 1951 CIA report on the implications of an agreement to limit all armed forces clearly identifies these concerns: “The more frequent, thorough, and unrestricted the inspection the less would be the possibility of Soviet concealment, but the greater would be the risk of sensitive disclosures by the US” (Central Intelligence Agency, 1951, 11). The report analyzes which side would gain more military advantage through “collateral” information gleaned by inspections. It recommends limiting inspections in order to protect US information even though inspections are the only way to assure Soviet compliance. Internal Soviet discussions are not readily available, but occasional public statements also reveal an appreciation of the tradeoff. In 1959 the USSR submitted a proposal on disarmament to the UN General Assembly, which sought to solve the “contradiction between the need to offer foreign inspectors access to military and other facilities and the apprehension that, so long as disarmament was not entirely completed, the information obtained during such inspections might be used to the detriment of national security” (Kokeyev and Androsov, 1990, 6).

Through most of the Cold War, the US generally demanded cooperative monitoring and the Soviets usually refused to accept any intrusive measures. There are important indications that this observed behavior was not the product of habit or ideological commitment, but rather explicit, though different, assessments of the transparency-safety tradeoff. In analyzing the Soviet position on monitoring in 1963, US analysts noted that the primary Soviet

concern was military vulnerability, as “territorial access might yield additional target information and presumably improve capabilities for a counter-force strike” ([Arms Control and Disarmament Agency, 1963](#)). US policymakers, on the other hand, were primarily concerned with transparency—would the US detect Soviet cheating on an agreement? However, some internal discussions are indicative of ongoing concerns about safety as well. For example, in discussing proposed inspections for the Chemical Weapons Convention, the Office of the Secretary of Defense supported a proposal for “unimpeded [inspections] access to all military or government-owned facilities,” while the intelligence community, Joint Chiefs, and State Department opposed “an arrangement by which the US would be obligated by treaty to provide the Soviet Union, or any other state, access to our sensitive, non-CW related military or intelligence facilities” ([Central Intelligence Agency, 1984](#)).

While the severity of the transparency-safety tradeoff typically ruled out open deals, closed deals were possible only when unilateral monitoring was sufficiently effective. The 1964 “Freeze” proposal called for limits on levels of intercontinental and medium-range ballistic missiles, bombers, and missile defense systems, as well as extensive on-site inspections to guarantee compliance. The USSR flatly rejected inspections. Internally, US policymakers also questioned the provisions which US negotiators were then advocating publicly: inspections to verify a ban on re-entry vehicles “might disclose sensitive US information on warhead or penetration aid characteristics” ([Arms Control and Disarmament Agency, 1964](#)). They concluded that inspections would not be in the US interest, but also that without additional observation it would be impossible to assess Soviet compliance, ruling out any deal.

Subsequent developments in satellite technology improved the efficacy of unilateral monitoring, so that each side could observe key elements of the other’s arming without inspections. The result was the 1979 SALT I agreement, which differed from the earlier Freeze proposal principally in that monitoring was downgraded from onsite inspections to satellite surveillance ([Parados, 1999, 22](#)). The US dropped its insistence on inspections, a key source

of Soviet opposition to an agreement (Bunn, 1992, 107). Revealingly, Freeze and SALT I also differed in precisely what was to be limited. US satellites could reliably observe Soviet launchers, but not the missiles themselves, and so negotiations shifted to limiting launchers. The SALT I proposals placed limits on silos and submarines, not the missiles themselves as had been the focus of the Freeze proposals (Graybeal and Krepon, 1988, 96). The ability to unilaterally monitor treaty compliance from space rendered a closed deal viable, but it also circumscribed the types of arms which could be controlled.

The INF Treaty, committing the superpowers to eliminate all land-based intermediate- and shorter-range missiles, was an unusual instance of an open deal, and was agreed only once the superpowers found a way to dampen the transparency-safety tradeoff. These missiles were mobile and therefore difficult to detect by satellite surveillance alone, so that unilateral monitoring would not suffice to assure compliance with their elimination. Inspections could ensure that existing missiles were eliminated and no new ones were produced, but they also posed serious risk of exposing other military secrets. A particular worry was that inspections of missile production facilities might reveal the technical details or numbers of missiles not slated for elimination, as well as information about other advanced military industry. US stealth technology, advanced radar, and industrial processes would have been of interest for Soviet operatives (Toth, 1988). During the negotiations, even nongovernmental experts traditionally supportive of arms control raised concerns that inspections would mean “Soviets crawling around our most sensitive production facilities” and leave the US at a relative disadvantage (Grier, March 11, 1987).

No treaty was agreed until creative ways were found to moderate this tradeoff. A notable example was in the monitoring of Soviet missile production. The Soviet SS-25 missile, which was not banned by the treaty and would stay in production, was similar to the banned SS-20 and manufactured in the same facility. The US needed to verify that this facility was not manufacturing banned missiles, but detailed inspection of all missiles leaving the facility

would reveal far more information about the SS-25 than the USSR was willing to accept (and more than the US would want to accept at its own facility being inspected by the USSR.) This issue was a potential deal-breaker for the treaty (Pifer et al., 2012, 12). It was solved by the US designing a specialized x-ray imaging machine that collected enough information to discriminate between the two missiles as they left the factory, but provided no additional information. The treaty was agreed, and this solution was implemented in inspections.¹⁹

Both the US and USSR saw the value of limiting their arming. But even though the term “arms control” has become nearly synonymous with their negotiations, the superpowers signed relatively few agreements and engaged in a terrifying, exorbitantly costly arms race. Their perception of a severe transparency-safety tradeoff explains why agreements emerged only when some arms became unilaterally observable, or when innovative technical solutions dampened the tradeoff in cooperative monitoring.

4.3 Earliest Attempts: The Interwar Arms Negotiations

Formal agreements to limit arming are relatively recent historically, beginning in the early 1700s and with most deals occurring after World War II. Among the earlier agreements were those made between the two world wars in an attempt to limit arms competition among the great powers, principally the US, UK, France, Germany, and Japan. The transparency-safety tradeoff helps explain why these agreements contained few measures for monitoring and why their limits dealt only with naval forces, and in particular with capital ships. Compliance with limits on large warships was assessed by states as easy to observe unilaterally, while inspections threatened to reveal military vulnerabilities, such as weaknesses in British capabilities. Ideas for limiting air or land forces faced the challenge of differentiating between military and commercial applications, for which a solution was not available.

¹⁹Soviet sensitivity to revealing secret information is evident in later accusations that this “CargoScan” device collected images that were several centimeters larger than allowed (Russell, 2001).

The interwar naval arms race was costly in two respects: quantitative increases in arms meant that more resources had to be spent on shipbuilding; and qualitative improvements meant that investment in innovation was needed to avoid the risk of strategic surprise by an adversary. Following World War I, Britain's naval power was challenged by the naval competition between Japan and the US, and Britain saw arms control as a way of maintaining its superiority while holding down naval spending (Maurer, 1994, 268–274). The Washington Naval Treaty and follow-on agreements set tonnage limits based on relative parity for Britain, the US, Japan, France, and Italy, which restricted the numbers of ships each could build. During WWI, states had also sought to build ever-larger “Dreadnought” battleships, leading to intense competition. The Chief of the British Naval Staff noted that the real problem for security and finance was “the principle of going one better” in naval shipbuilding because it allowed one state to surprise another with design-related advancements that could not be met as easily as building extra ships to match numbers (CAB 29/148, NCM (35) 23, 30 Oct. 1934, p. 2). The international agreements set a qualitative restraint on this race by limiting sizes of ships and their guns. The naval treaties restrained arming competition and shaped state choices about naval construction until the mid 1930's, after which point political changes in Germany and Japan made further limitations untenable, as neither of these states stuck to prior commitments (Maurer, 1994, 289).

The Washington Naval Treaty called for states to declare the sizes and characteristics of ships being constructed, but did not include provisions for cooperative monitoring. States' assessment of the relative value of transparency in an agreement compared to the risks posed to safety of military information helps explain why the Washington Treaty and follow on negotiations call for little monitoring, or “supervision” and “investigation” as phrased at the time. The signed agreements are examples of open deals at acceptable only at a very low monitoring level. Agreements over naval forces stand in contrast with failed negotiations during this period – including lack of limits on arms racing in air and land forces – where dual

use technology made observation difficult and exacerbated the transparency-safety tradeoff.

France was the key proponent of intrusive inspections, primarily due to concerns over German arming (Kitching, 2003; Burns, 1969, 97). However, most of the other states, and particularly Britain, strongly opposed any kind of intrusive monitoring. As a result, limitations which would require such measures were not considered at the conference (Kitching, 2003). Most of the parties—apart from France—believed that greater transparency had limited benefits. Battleships could be monitored through espionage and even publicly available information such as media reports. Capital ships were massive in size and typically built in the open over the course of around two years, leaving plenty of time for detection.²⁰ The technology for building dreadnoughts had been fairly transparent, as evidenced by the quick diffusion of these capabilities across states. The technology was also relatively mature and new developments came slowly, so states did not expect a situation where the lack of close monitoring would miss a significant military innovation (Kearn Jr, 2014, 76). Inspection was seen as having some marginal advantages; observing exact differences in tonnage for ships was difficult without more intrusive access, so it was known that states could cheat by building ships with stronger metal protection (Maiolo, 1999, 108).

At the same time, the UK saw openness to monitoring as jeopardizing the safety of other sensitive military information, namely the reality of its military weakness and practical inability to honor defensive treaty commitments. The UK rejected proposals that included “continuous and automatic” verification (Kitching, 2003, 160–161). Internally, officials expressed concern that “supervision” of limitations would “lead to the disclosure of [UK] war reserves and stocks” (CAB 27/509, DC(M) (32) 15, 24 Aug. 1932, p. 2). This information would be detrimental to British security because “[war material] stocks have dwindled to a

²⁰Maiolo (2016, 62) notes that, in contrast to the quick expansion of armies and air forces, surface naval vessels like battleships, aircraft carriers, and cruisers took three to six years to build and bring into service. These large warships were difficult to conceal and, thus, compliance with treaty obligations was easy to verify (Overy (2016, 129–130); Kaufman (1990, 103)).

dangerously low level, acceptance of which is entirely dependent upon the facts remaining unknown to other countries” (CAB 27/509, DC(M) (32) 15, 24 Aug. 1932, p. 3). In a cabinet meeting with the prime minister, First Lord of the Admiralty Sir Bolton Eyres-Monsell explicitly stated that, “objections to Supervision were not because we had anything to hide but because of the bareness of the cupboard.”²¹ The British Cabinet concluded that increased transparency would increase insecurity by revealing that Britain would be unable to honor its commitment to intervene on France’s behalf in the event of a conflict; “to accept a form of permanent and rigorous inspection would surely only to show France that our obligation under Locarno was, in fact, of no use at all — in other words, such a form of supervision, by disclosing the nakedness of the land, would tend to counteract any value which the Locarno Treaty might have to France at the present time.”²²

The US had also assessed a high cost for the safety side of the tradeoff. US Navy feared that Japan would have more to gain from inspections than the US (Burns, 1969). During the Disarmament Conference, the Navy worried that inspections would reveal too much to the Japanese about US defenses in the Philippines, a concern which outweighed warnings that the Japanese may cheat on the agreement²³ (Kaufman, 1990, 104).

The Washington Naval Treaty’s focus on capital ships is also revealing of how the tradeoff can limit the extent of acceptable limitations in a multi-faceted arms race. In contrast to limits on ships, the negotiating states could not agree on how to address limitations on air and land forces (Roskill, 1969, 322–323). In these cases, the transparency-safety tradeoff was even more severe, and monitoring limitations would mean significant costs to military safety.²⁴

²¹In a statement that bears striking similarity to Cold War era concerns over Russian inspectors, Eyres-Monsell also questioned whether the US, despite President Roosevelt’s stated support for supervision, would actually, “agree to have Japanese sitting down in Washington, with right to inspect anything they liked in America[.]” (CAB 27/505, DC(M) (32) 17th Cons., 20 June 1933, pp. 3–4).

²²Id., at p. 8, also cited in Kitching (2003, 172).

²³Notably, top US political leaders had different views. President Franklin D. Roosevelt and Secretary of War Henry Stimson supported inspections. In 1932-1933 the official US position shifted to supporting some inspections, though did not go as far as the French (Burns, 1969).

²⁴The transparency-safety tradeoff was one of several factors at work in this case. The lack of agreement

Aviation and land forces, built on the new technology of the combustion engine, were seen as very difficult to monitor. This was primarily because the relevant technologies were dual use; applications were widespread for both for civilian and military purposes (Kearn Jr, 2014). In principle, states could differentiate between civilian and military applications though direct monitoring of factories, bases, and other military installations. Indeed this high level of inspection was used for monitoring German disarmament in the Versailles treaty. However, for most of the negotiating parties, using such intrusive tools to monitor one another was not seen as an option. Officials at the time explicitly cited dual use applications as the reason for their opposition to limits.²⁵

The interwar arms treaties are often seen as a failure: they prevented neither an arms race nor another world war among the great powers. However, they did restrain at least some degree of arming by the major powers for over a decade: less money was spent on battleships even as other forms of arming continued. The transparency-safety tradeoff explains why grand ideas for more comprehensive arms limitations were never implemented.

5 Conclusion

Just as with war, the costs of arming imply that agreements to avoid it should be preferable for states, and yet arming is ubiquitous and its control very rare. We have argued that an important obstacle to arms control is the tradeoff involved in monitoring: transparency is required to assure one side of the other's compliance with arms limits, but might also reveal vulnerabilities that could be exploited by the first side in an arms race or war. If this

on air and land forces also had to do with differences in assessment over the relative threat of naval versus other forces: air power was recognized as a threat to battleships, but both American and British officials believed the capital ships were the backbone of naval fleets (Roskill, 1969, 322).

²⁵For example US Secretary of State Charles Hughes opposed limits on military aircraft for this reason (Kaufman, 1990, 56). The British were interested in ideas on abolishing naval and military aircraft, or a ban on "bombing from the air" (ADM 116/2999, March 1934, p. 19 (Part V - Aircraft Carriers, ¶49)), but likewise assessed dual use applications as a key challenge for limiting aviation.

transparency-safety tradeoff is severe, then any cooperative monitoring that would assure one side would be unsafe for the other, and so no cooperation in monitoring is possible. If unilateral monitoring does not render a side's arming transparent enough, then no arms limits will be agreed. We have shown, across a diverse range of contexts, that this tradeoff helps to explain whether arms control occurs, what precise limits are agreed, and what monitoring is accepted.

Our theory identifies at least three factors that should alleviate the transparency-safety tradeoff and thereby improve the prospects for arms control, yielding observable implications which might be tested in future work. First, states can resort to a third-party monitor. The monitor could be given the necessary access to a state's territory in exchange for revealing no information beyond whether that state was complying with an arms control deal. In principle, this averts the tradeoff altogether, since both transparency and safety can be achieved. However, such a third-party can only be effective if both sides trust that it is able and willing to uphold its duty. One state may fear that the third-party is not impartial, so that it could not be counted upon to reveal discovered evidence of cheating. What third-party would either superpower have trusted to keep it safe from covert arming by the other superpower? Alternatively, a state may fear that its adversary would penetrate the third-party's monitoring, and thereby discover information that was not supposed to be revealed, as the US reportedly did with UNSCOM in Iraq.

Second, advances in sensing technology can increase the efficacy of unilateral monitoring, so that cooperative monitoring is not required for arms control to be viable. The number of arms control agreements made has increased exponentially over time, with most occurring in the latter half of the twentieth century (Vaynman, 2017). We conjecture that this increase is related to the rise of radio communication, overhead imagery, and now the internet, all of which provided new avenues for unilateral monitoring of arms control deals. Clearly, spy satellites and the unilateral monitoring they enabled were key to making some of the Cold

War deals we examined.

Third, the compliance with arms control of states that feature generally more open societies should be easier to monitor unilaterally. Thus, such states should find it easier to make an arms control deal with each other than with less open states. We found that the US believed its general openness left it at a disadvantage to the more closed USSR, leading the US to favor cooperative monitoring and the USSR to resist it. Generalizing, we conjecture that more open dyads should be more likely to make arms control agreements than more less open dyads, conditional on the members of the dyad being adversaries.

A better understanding of the transparency-safety tradeoff might also inform assessments of new arms control deals. Consider the recent deal between the US and Iran on limiting the latter's nuclear program. Our theory helps to explain why it features quite comprehensive monitoring of non-military facilities such as nuclear reactors and centrifuge plants, but no monitoring of military facilities such as the bases at which high-explosives testing—essential to building nuclear weapons—was reportedly carried out. Cooperative monitoring of the former poses a low risk to Iran of revealing other security-relevant information, but giving international inspectors easy access to some military bases might enable the US to learn much more than simply whether Iran was complying with the deal.²⁶ Because of the restriction to non-military facilities, the agreed monitoring will not provide perfect transparency into all aspects of Iran's program. However, criticizing the deal for this reason fails to recognize that a deal with perfect transparency may not be viable, because it would be too unsafe for Iran, and that even imperfect transparency may be enough to support a deal. Those who wish to see monitoring improved in future negotiations might do better to devise monitoring arrangements that would dampen or avert the transparency-safety tradeoff, so that improvements would be acceptable to both sides.

²⁶The tradeoff is not completely absent even in the monitoring arrangements that were made. Long (April 3, 2015) argues that these will provide targeting intelligence that should make it easier for the US to attack Iran's program in the future.

References

- ADM 116/2999. March 1934. “Memorandum by the Chief of the Naval Staff in Preparation for the 1935 Naval Conference.” The National Archives of the UK (TNA).
- Arms Control and Disarmament Agency. 1963. “Meeting on Arms Control Issues.” Office of Politico-Military Affairs – Records Relating to Disarmement and Arms Control, 1961–66.
- Arms Control and Disarmament Agency. 1964. “Meeting of the Committee of Principals.” Office of Politico-Military Affairs – Records Relating to Disarmement and Arms Control, 1961–66.
- Baliga, Sandeep and Tomas Sjöström. 2008. “Strategic Ambiguity and Arms Proliferation.” *Journal of Political Economy* 116(6):1023–1057.
- Bas, Muhammet A. and Andrew J. Coe. 2012. “Arms Diffusion and War.” *Journal of Conflict Resolution* 56(4):651–674.
- Bas, Muhammet A. and Andrew J. Coe. 2016a. “A Dynamic Theory of Nuclear Proliferation and Preventive War.” *International Organization* 70(4):655–685.
- Bas, Muhammet A. and Andrew J. Coe. 2016b. “Give Peace a (Second) Chance: A Theory of Nonproliferation Deals.” Paper presented at the 2016 Annual Meetings of the American Political Science Association.
- Benson, Brett and Quan Wen. 2011. A Bargaining Model of Nuclear Weapons Development and Disarmament. In *Causes and Consequences of Nuclear Proliferation*, ed. Robert Rauchhaus, Matthew Kroenig and Erik Gartzke. Taylor and Francis pp. 45–62.
- Bunn, George. 1992. *Arms Control by Committee: Managing Negotiations with the Russians*. Stanford University Press.

- Burns, Richard Dean. 1969. "International Arms Inspection Policies Between the World Wars, 1919–1934." *The Historian* 31(4):583–603.
- CAB 27/505, DC(M) (32) 17th Cons. 20 June 1933. "Conclusions of a Meeting of the Committee held in the Prime Minister's Room, House of Commons, on Monday, 19th June, 1933 at 6.30 p.m." The National Archives of the UK (TNA).
- CAB 27/509, DC(M) (32) 15. 24 Aug. 1932. "Memorandum by the Admiralty on Subjects other than Direct Naval Limitation." The National Archives of the UK (TNA).
- CAB 29/148, NCM (35) 23. 30 Oct. 1934. "Memorandum by the Chief of Naval Staff." The National Archives of the UK (TNA).
- Carnegie, Allison and Austin Carson. 2017a. "The Disclosure Dilemma: International Organizations, Intelligence, and Sources and Methods." Working Paper.
- Carnegie, Allison and Austin Carson. 2017b. "The Spotlight's Harsh Glare: Rethinking Publicity and International Order." Working Paper.
- Central Intelligence Agency. 1951. "Special Estimate: Intelligence Implications of a Census and Verification of Armed Forces and Armaments." National Intelligence Council (NIC) Collection.
URL: <https://www.cia.gov/library/readingroom/document/0000170378>
- Central Intelligence Agency. 1984. "Memorandum for Director of Central Intelligence." General CIA Records.
- Coe, Andrew J. 2017. "Containing Rogues: A Theory of Asymmetric Arming." Working Paper.
- Coe, Andrew J. and Jane Vaynman. 2015. "Collusion and the Nuclear Nonproliferation Regime." *Journal of Politics* 77(4):983–997.

- Conrad, Justin, Hong-Cheol Kim and Mark Souva. 2013. "Narrow Interests and Military Resource Allocation in Autocratic Regimes." *Journal of Peace Research* 50(6):737–750.
- Dai, Xinyuan. 2002. "Information Systems in Treaty Regimes." *World Politics* 54(4):405–436.
- Dai, Xinyuan. 2007. *International Institutions and National Policies*. Cambridge University Press.
- Debs, Alexandre and Nuno P. Monteiro. 2014. "Known Unknowns: Power Shifts, Uncertainty, and War." *International Organization* 68(1):1–31.
- Department of State. 1999. *Saddam Hussein's Iraq*. United States Government.
- DiGiuseppe, Matthew. 2015. "Guns, Butter, and Debt: Sovereign Creditworthiness and Military Expenditure." *Journal of Peace Research* 52(5):680–693.
- Downs, George W. and David M. Rocke. 1990. *Tacit Bargaining, Arms Races, and Arms Control*. University of Michigan Press.
- Dunn, Lewis A. 1990. "Arms Control Verification: Living with Uncertainty." *International Security* 14(4):165–175.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.
- Fearon, James D. 2011. "Arming and Arms Races." Paper presented at the 2010 Annual Meetings of the American Political Science Association, Washington, DC.
- Fearon, James D. 2016. "Cooperation and Conflict under Anarchy." Working Paper.

- Fordham, Benjamin O. and Thomas C. Walker. 2005. "Kantian Liberalism, Regime Type, and Military Resource Allocation: Do Democracies Spend Less?" *International Studies Quarterly* 49(1):141–157.
- Gallagher, Nancy W. 1997. "The Politics of Verification: Why 'How Much?' Is Not Enough." *Contemporary Security Policy* 18(2):138–170.
- Gallagher, Nancy W. 1999. *The Politics of Verification*. Johns Hopkins University Press.
- Goldblat, Jozef. 2002. *Arms Control: The New Guide to Negotiations and Agreements with New CD-ROM Supplement*. Sage.
- Goldsmith, Benjamin E. 2003. "Bearing the Defense Burden, 1886–1989: Why Spend More?" *Journal of Conflict Resolution* 47(5):551–573.
- Graybeal, Sidney N. and Michael Krepon. 1988. On-site Inspections. In *Verification and Compliance: A Problem-Solving Approach*, ed. Michael Krepon and Mary Umberger. London, UK: The Macmillan Press.
- Grier, Peter. March 11, 1987. "Arms treaty could let 'sanctioned spies see US weapons plants." *Christian Science Monitor* .
- Harvey, Frank P. 2011. *Explaining the Iraq War: Counterfactual Theory, Logic and Evidence*. Cambridge University Press.
- Iraq Intelligence Commission. 2005. "Report of the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction." Unclassified Version, March 31.
- Iraq Survey Group. 2004. "Comprehensive Report of the Special Advisor to the DCI on Iraq's WMD." September 30.

- Jackson, Matthew O. and Massimo Morelli. 2009. "Strategic Militarization, Deterrence and Wars." *Quarterly Journal of Political Science* 4(4):279–313.
- Kaufman, Robert Gordon. 1990. *Arms Control During the Pre-nuclear Era*. New York, NY (USA): Columbia University Press.
- Kearn Jr, David W. 2014. *Great Power Security Cooperation: Arms Control and the Challenge of Technological Change*. Lexington Books.
- Kilgour, D Marc. 1994. "The Use of Costless Inspection in Enforcement." *Theory and Decision* 36(3):207–232.
- Kitching, Carolyn J. 2003. *Britain and the Problem of International Disarmament: 1919–1934*. Routledge.
- Koblentz, Gregory D. Forthcoming. "Saddam versus the Inspectors: The Impact of Regime Security on the Verification of Iraq's WMD Disarmament." *Journal of Strategic Studies* .
- Kokeyev, Mikhail and Andreï Androsov. 1990. *Verification—The Soviet Stance: Its Past, Present, and Future*. New York: UNIDIR.
- Koremenos, Barbara. 2001. "Loosening the Ties that Bind: A Learning Model of Agreement Flexibility." *International Organization* 55(2):289–325.
- Krass, Allan S. 1985. "The Politics of Verification." *World Policy Journal* 2(4):731–752.
- Kreps, Sarah E. 2016. "The Institutional Design of Arms Control Agreements." *Foreign Policy Analysis* .
- Kydd, Andrew. 1997. "Game Theory and the Spiral Model." *World Politics* 49(3):371–400.
- Kydd, Andrew. 2000. "Arms Races and Arms Control: Modeling the Hawk Perspective." *American Journal of Political Science* 44(2):228–244.

- Lindsey, David. 2015. "Military Strategy, Private Information, and War." *International Studies Quarterly* 59(4):629–640.
- Long, Austin. April 3, 2015. "If You Really Want to Bomb Iran, Take the Deal."
- Maiolo, Joseph. 2016. Between the Two World Wars: Introduction. In *Arms Races in International Politics: From the Nineteenth to the Twenty-First Century*, ed. Thomas G Mahnken, Joseph Maiolo and David Stevenson. Oxford University Press.
- Maiolo, Joseph A. 1999. "The Admiralty and the Anglo-German Naval Agreement of 18 June 1935." *Diplomacy and Statecraft* 10(1):87–126.
- Maurer, John H. 1994. Arms Control and the Washington Conference. In *The Washington Conference, 1921–1922: Naval Rivalry, East Asian Stability and the Road to Pearl Harbor*, ed. Erik Goldstein and John Maurer. Frank Cass pp. 267–293.
- Meirowitz, Adam and Anne E. Sartori. 2008. "Strategic Uncertainty As a Cause of War." *Quarterly Journal of Political Science* 3(4):327–352.
- Nordhaus, William, John R. O Neal and Bruce Russett. 2012. "The Effects of the International Security Environment on National Military Expenditures: A Multicountry Study." *International Organization* 66(3):491–513.
- Overy, Richard. 2016. Aircraft and the Arms Race Between the World Wars. In *Arms Races in International Politics: From the Nineteenth to the Twenty-First Century*, ed. Thomas G Mahnken, Joseph Maiolo and David Stevenson. Oxford University Press.
- Parados, John. 1999. Prague Spring and SALT: Arms Limitation Setbacks in 1968. In *The Foreign Policies of Lyndon Johnson: Beyond Vietnam*, ed. H.W. Brands. College Station, Texas: Texas A&M University Press.

- Pifer, Steven, Avis T Bohlen, William F Burns and John Woodworth. 2012. *The Treaty on Intermediate-Range Nuclear Forces: History and Lessons Learned*. Washington, DC: Brookings Institution Press.
- Powell, Robert. 1993. "Guns, Butter, and Anarchy." *American Political Science Review* 87(1):115–132.
- Roskill, Stephen Wentworth. 1969. *Naval Policy Between the Wars: The Period of Anglo-American Antagonism 1919–1929*. Vol. 1 Walker & Co.
- Russell, John. 2001. "On-Site Inspections Under the INF Treaty: A Post-Mortem." London, UK: VERTIC Breifing Paper 01/02.
- Toth, Robert C. 1988. "Verification of Arms Pact May Open Spy Door." *Los Angeles Times* .
- Verdier, Daniel. 2008. "Multilateralism, Bilateralism, and Exclusion in the Nuclear Proliferation Regime." *International Organization* 62(3):439–476.
- Wittman, Donald. 1989. "Arms Control Verification and Other Games Involving Imperfect Detection." *American Political Science Review* 83(3):923–945.