Why Arms Control Is So Rare

Andrew J. Coe and Jane Vaynman*

April 21, 2018

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 great power competition in arms in the interwar period; and the superpowers’ military rivalry during the Cold War. These arms races account for almost 40% of all global arming in the past two centuries.

Word Count: 11,828

*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.
Why do states spend so much on arms? About 2.2% of the world’s wealth—$2.7 trillion—was devoted to its militaries in 2016.\footnote{This is the most recent figure from the CIA World Factbook.} How much is a mere 2.2 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.\footnote{Of course, not all military spending is lost to the civilian economy. Even if only 50% is, the remainder would still suffice to pay for all these things, as we document in the online appendix.}

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?\footnote{See Fearon (1995) for the seminal explication of this inefficiency puzzle of war.} 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, Forthcoming; Benson and Wen 2011; Coe Forthcoming; Debs and Monteiro 2014; Fearon 2011, Forthcoming; Jackson and Morelli 2009; Meirowitz and Sartori 2008; Powell 1993).\footnote{Fearon (2011, Forthcoming); 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. Arming might also arise from asymmetric information about preferences rather than actions, as in theories of the security dilemma. However, if arming is perfectly observable, it should be possible to make an arms control deal regardless of whether a state is known to be “status quo-oriented” or “revisionist.” See Kydd (1997, 2000) for reviews and formalizations of this earlier literature.}
This certainly accords with 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.

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. 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

\(^5\)The only exceptions are Bas and Coe (Forthcoming); Coe (Forthcoming), but both models still assume that arming is imperfectly observable, so that a deal to limit arming is often not viable.
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 show that this tradeoff bedeviled attempts between the world wars 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.

Finally, 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. 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 compromises to intrusiveness were made to ensure that inspectors would be able to see only what was necessary to verify the limit, with no other security-relevant information exposed.
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 Forthcoming), 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 (Forthcoming, 2017), 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) and Slantchev (2010) show 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 recognized that worries about espionage sometimes impeded negotiations over verification, but instead viewed domestic politics as the most critical determinant of arms control (Dunn 1990; Gallagher 1997, 1999; Goldblat 2002; Krass 1985).

The few empirical analyses of military spending 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). These studies analyze cross-national differences in these expenditures but do not explain why they are almost always positive and large.
Model Setup

A (the monitoring state, referred to as feminine) and B (the arming state, masculine) bargain over revisions to a prior settlement of a set 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.\(^6\) 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.\(^7\) 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 \(\lambda\) 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

---

\(^6\)The 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.

\(^7\)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.\footnote{In 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.} If \( B \) is closed to monitoring and has invested, then with probability \( \tau_c \) 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.\footnote{Allowing 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.} 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 \( W^A_n \) and \( W^B_n \) respectively. If \( B \) is not newly-armed, the values are \( W^A_o \) and \( W^B_o \) if he is open to monitoring and \( W^A_c \) and \( W^B_c \) if he is not. We assume \( W^A_o \geq W^A_c > W^A_n \) and \( W^B_n > W^B_c \geq W^B_o \). It will sometimes be convenient to treat \( W^A_o \) and \( W^B_o \) as functions of \( \tau_o \), with \( W^A_o(\tau_c) = W^A_c \) and \( W^B_o(\tau_c) = W^B_c \). We assume that \( W^A_o(\cdot) \) is nondecreasing and \( W^B_o(\cdot) \) 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.\footnote{This is necessary only to avoid a surfeit of quantities like \( \max \{ W^B_o, 0 \} \) in the exposition.} Conflict is costly: the value destroyed by it is \( D^w_x \equiv \frac{1}{1-\delta} - W^A_x - W^B_x > 0 \), for any \( x \in \{ n, o, c \} \), with \( D^w_o \leq D^w_c \).

\( A \)'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-
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 B is closed, then the A’s
ability to observe arming is limited to unilateral intelligence-gathering. If B is open, then
this ability is supplemented by bilateral cooperation. Cooperation 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 ($W^B_o \leq W^B_c$ and
$W^A_o \geq W^A_c$). Arming openly may also reduce the costs of this outcome, as it makes possible
a more precise reaction by A to B’s arming ($D^w_o \leq D^w_c$). 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 $\tau_c$ or $\tau_o$). 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 $\tau_o$. However, allowing this bargaining to occur would not change our key finding that, under certain conditions, a deal would not be viable for *any* degree of openness. The model presented here also assumes that only one side can arm and be monitored. In the online appendix, we show that qualitatively similar results hold when both sides can arm and be monitored.

In other respects, our setup is similar to that of models of nuclear proliferation, such as Bas and Coe (2016, Forthcoming); Benson and Wen (2011); Coe (Forthcoming); 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 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 arming 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 expect similar results could be obtained using the latter type.

**Analysis**

We first establish the “no-deal” equilibrium, governing behavior in the absence of a deal to stop arming. This forms the backdrop for our analysis of the possibilities for arms control deals and how these are affected by the transparency-safety tradeoff.\(^{11}\) 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.

---

\(^{11}\) 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 this entails a deal being made.
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 no need for any deal. We discard this uninteresting case.

**Assumption 1.** *In the absence of a deal, $B$ would invest, given the chance: $k < \delta \lambda [W^B_n - W^B_c]$.***

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 $W^A_c + \overline{V}^B_c > \frac{1}{1-\delta} - \frac{k}{1-\delta(1-\lambda)}$, where $\overline{V}^B_c \equiv \max \left\{ W^B_c, \frac{-k + \delta \lambda W^B_n}{1-\delta(1-\lambda)} \right\}$.***

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

---

12Here 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.
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 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 remain closed to monitoring has nothing to do with concealing 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 \((W^B_c \geq W^B_o)\), 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 \(V^A_c\) and \(V^B_c\)). 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 \(V^A_p \equiv V^A_c\), \(V^B_p \equiv V^B_c\), and \(S \equiv \frac{1}{1-\delta} - V^A_p - V^B_p\). There is a deal in which B is closed to monitoring if and only if \(\delta [\lambda + \tau c (1 - \lambda)] S \geq -k + \delta \lambda \left[ W^B_n - V^B_p \right] \).

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 ($\delta \lambda [W^B_n - V^B_P]$) 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 ($\lambda$) 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 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.

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 ($\tau$) 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 ($\tau_c$) 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. 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 reneging 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 $\tau_c$ is replaced by $\tau_o$ 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 ($\tau_o$) 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 ($W_A^c = W_A^o$ and $W_B^c = W_B^o$). 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 ($\tau \leq \tau_o < 1$).*

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 needed.

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. Let $\tau_o$ be the largest value of $\tau_o$ such that $W^A_o(\tau_o) - k + \lambda W^B_n + (1 - \lambda)V_c^B \leq \frac{1}{1 - \delta}$.  

14
Corollary 3. An open deal is possible if and only if the transparency-safety tradeoff is not too severe ($\tau \leq \tau_o$).

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: $\tau_o$ must be at least $\tau$. 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”: $\tau_o$ can be no more than $\tau_o$.

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 ($[\tau, \tau_o]$) 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.
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 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 and unilateral monitoring because of a severe transparency-safety tradeoff. Finally, we examine the negotiations between the US and USSR over limiting the Cold War arms race, finding that our tradeoff was a constant concern for both sides.

The cases illustrate how the tradeoff shapes whether an agreement is made, what kind of monitoring it features, and what is limited under it. However, they are not merely illustrative: they occurred during arms races which account for almost 40% of the world’s military spending for the last two centuries, or equivalently 40% of failures to limit arms. In particular, the main parties to the interwar arms talks account for over half of all military spending from 1919 to 1938, while the superpowers made up almost two-thirds of spending from 1946 to 1991. Understanding arms control and its absence during these periods is thus important in its own right.

13 All figures are computed from the National Military Capabilities dataset and adjusted for inflation. Spending by the belligerents during the world wars is excluded from the total since total warfare, rather than our tradeoff, ruled out arms control deals.

The United States imposed sanctions on, undertook limited strikes against, and eventually invaded Iraq to overthrow its regime in 2003, all largely for the purpose of ending its pursuit of WMD. We now know that Iraq had no WMD and suspended its programs by 1996, but 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 the resulting conflict were tremendous for both the US and Iraq. Both 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 (2018), combining these with our own arguments about unilateral monitoring to explain why no deal was viable.\footnote{The 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 (2018) argues that Saddam had incentives to lie about the true cause of his 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.}

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 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 (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 disarmament could not be assured otherwise.

Thus, only cooperative monitoring could provide the transparency required 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 Iraq War assume that this monitoring was necessarily flawed (Coe Forthcoming; 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, Iraq could be blanketed with 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). 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 2018, 23–24).15 These efforts and those of independent elements within Iraq resulted in a number of popular uprisings, coup plots, and assassination

15We henceforth use only page numbers in citing this article.
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 (Department of State 1999). 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 concealing 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 (16–17). It seems very unlikely the US would not have used inspections of those sites 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 Saddam’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, during this period, the US constantly demanded more
transparency, Iraq resisted in order to protect the regime’s safety (4), and no deal was made.

Closed Deal: The Interwar Arms Negotiations

Formal agreements to limit arming are relatively recent historically, with very few occurring
before the 20th century and most agreed after World War II (Vaynman 2017). 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 Washington and London Naval Treaties set limits on the size, total tonnage, and gun
caliber of the participants’ large warships, suppressing naval arms competition until the
mid-1930’s. The transparency-safety tradeoff helps explain why these agreements relied
upon unilateral monitoring and why their limits dealt only with large warships. Compliance
with these limits was perceived to be easy to monitor unilaterally, making a closed deal
possible. By contrast, verifying limits on other forces would require inspections. Because
these inspections risked revealing important military secrets, such as concealed weaknesses
in British capabilities, they imposed a severe transparency-safety tradeoff, preventing any
agreement to limit those forces.

Limits on capital ships could be monitored quite easily through unilateral means such
as espionage and media reports. These ships were massive and typically built in the open,
so that spies could readily determine whether construction was ongoing (Overy 2016, 129–
130). Moreover, they required years to complete, leaving plenty of time for detection and

\[16\] Subsequent political changes in Germany and Japan led them to abandon these treaties and made
further limitations untenable (Maurer 1994, 289).
further investigation (Maiolo 2016, 62). The British Admiralty assessed that “with regard to numbers of ships or the size of guns carried, investigation is clearly unnecessary, since there can be no real doubt as to the facts in such matters” (British Cabinet 1932, 4) and that “it would not be long before the Admiralty heard from one source or another of the construction of a battleship or a submarine” (DBFP 1968, 930). The US also “regarded the capital ship as easy to count and difficult to hide” (Kaufman 1990, 57). Indeed, the US began the Washington Naval Conference by “listing by name every ship in the three major navies that was suitable for scrapping” (Tooze 2014, 397), and British documents discussing the proposed limits had detailed information on all the involved ships (British Cabinet 1921, 59–68, 74). France too agreed that “it is rather difficult to lay the keel of a ship in the stocks, to prepare the dock yards, without the world knowing it” (British Cabinet 1921–22, 148). Because serious violations of agreed limits were difficult to conceal, unilateral monitoring sufficed to meet the transparency requirement, and deals to limit large warships were made.

The great powers also considered limiting other forces. Monitoring limits on military personnel posed challenges because conscription provided for rapid expansion (British Cabinet 1921–22, 148, 166). The British observed that “so long as Germany has any men of military age they must of course be regarded as potential soldiers and the German Staff will no doubt continue to work out plans for organising large numbers in formations and units” (British Cabinet 1921, 149). The problem was not limited to Germany: “conscription provides France with an unlimited number of ready-made mechanics of highest quality who can be drafted into air force from all trades in time of emergency, thus obviating peace expenditure on training and providing for immediate expansion of air force on mobilization” (British Cabinet 1921, 218, 244). The US “recognized at the outset that it would be difficult, if not impossible, to provide at this [Washington Naval] Conference for the limitation of land forces” (FRUS 1938, 314).

Limits on mechanized forces, whether on land or in the air, posed similar difficulties
for verification because commercial vehicle factories could be quickly turned to military production (British Cabinet 1921–22, 702). The British noted “the narrow line between commercial and military aviation” (DBFP 1966a, 475–476) and observed that air would be easier to limit if military aviation “developed on lines different from civil aviation” (British Cabinet 1921, 218). The problem “was to try to create a verification regime that would ensure that military aviation was not being sustained by some backstairs artifice [...] or to be certain that in the event of war civilian aircraft would not simply be converted into bombers” (Overy 2016, 116). “In the 1930s it was difficult to calculate the capacity of an enemy aircraft industry as it rapidly expanded; estimates of enemy air strength could be calculated from effective photo-reconnaissance or knowledge of the organizational structure of the force, but these would reveal little about the situation of reserves or the potential for wartime mobilization” (Overy 2016, 129–130). The US “ruled out any direct limits on military aircraft because nations could not verify where the civilian aircraft industry ended and military use began” (Kaufman 1990, 56).

Despite these challenges for monitoring, it was still possible to verify compliance with limits on land and air forces. This was accomplished in Germany under the Treaty of Versailles with quite intrusive “supervision” (British Cabinet 1932, 5). The Allied “Supreme Council came to the decision ‘that the Allied Governments shall assure themselves by constant supervision that Germany is fulfilling her obligations’ ” including in the need “to distinguish civil aviation from military aviation” (DBFP 1968, 927–928). The British believed this supervision worked, because as it was withdrawn in the 1920s, they worried whether they would still be able to observe continued German compliance (DBFP 1966b, 190–197). The French, despite being inclined to stringency toward Germany, admitted that this intrusive monitoring scheme had worked and thus was sufficient to verify compliance (British Cabinet 1921–22, 146–147). The British, despite being inclined to leniency, nevertheless believed such inspections remained necessary to monitor limits on German land forces (DBFP 1968,
Owing to its geographical situation, France was more concerned with air and especially land power than the other parties to the negotiations (Kitching (2003, 95), British Cabinet (1932, 9, 11), British Cabinet (1921–22, 94, 129–148, 151–157, 398–399). “The French Government maintain that the supervision of German aviation is a matter involving French national security and one on which they cannot compromise” (DBFP 1968, 901). France therefore required the greatest transparency, proposing “permanent, automatic, and continuous supervision” before agreeing to limits on these forces (Kitching 2003, 156, 162). British negotiators perceived that “France would undoubtedly weaken even on the question of the maintenance of her superiority in war material, provided the question of supervision and control can be settled to her satisfaction. If an effective and permanent organ of supervision can be agreed upon, then the danger of surprise and sudden aggression is much diminished” (British Cabinet 1932, 13).

By contrast, Britain had the most to lose from transparency and therefore consistently resisted any thorough system of inspection (British Cabinet 1932, 3). Britain’s military opposed inspections out of fear that these would reveal a lack of defense preparations and other secrets (Kitching (2003, 156, 160–161), British Cabinet (1932, 2–4). The Admiralty assessed that “Investigation might lead to disclosure of stocks of war material. In some respects these stocks have dwindled to a dangerously low level, acceptance of which is entirely dependent upon the facts remaining unknown to other countries. Again, disclosure of stocks of certain materials may disclose also certain portions of our defense preparations” (British Cabinet 1932, 3). If the “bareness of the cupboard,” as the First Lord of the Admiralty characterized it (British Cabinet 1933, 4), were revealed, France would have reason to doubt Britain’s guarantee of its security from German attack under the Treaty of Locarno, undermining the delicate system of alliances and guarantees undergirding peace ((Kitching 2003, 172), (British Cabinet 1932, 10–11). “For us to have to accept a form of permanent and rigorous
inspection would only tend to shew 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” (British Cabinet 1933, 9). As a result, Britain never agreed to the supervision France sought (British Cabinet 1933). For its part, the US Navy “thought that [inspection] visits ‘would lose us more by the information we might give out.’ ” (Kaufman 1990, 57). In particular, “the United States stood to lose more than it would gain by on-site inspections, lest the Japanese discover too much about the American defenses in the Philippines” (Kaufman 1990, 104).

These differences over monitoring scuttled the various attempts at general disarmament (British Cabinet 1932, 5–6), (British Cabinet 1921, 159–161), (Kitching 2003, 160–172). As a result, the interwar arms conferences are often seen as a failure: they prevented neither an arms race nor another world war. However, they did restrain some arming by the great powers for over a decade: less money was spent on large warships even as other forms of arming continued. The transparency-safety tradeoff explains why more comprehensive limits were not agreed.

Open Deal: Verification during the Cold War

The US and USSR undertook the most expensive arms race in human history. 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 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. Agreements
with provisions for cooperative monitoring did emerge when the states moderated the tradeoff using the design of the agreement itself, intentionally selecting lower levels of transparency in order to protect the safety of other military capabilities. Here, we summarize our findings across several negotiations during the Cold War. For brevity’s sake, documentation of these claims is largely relegated to the online appendix.

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 noted the problem of having inspections before general disarmament was completed; though needed, such inspections would also reveal information detrimental to security.

Through most of the Cold War, the US generally demanded cooperative monitoring and the USSR usually refused to accept any intrusive measures. The historical record provides 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. The USSR assessed particularly high risks in allowing for greater openness, while the US saw greater risks in the possibility of missing Soviet cheating on an agreement. In describing the Soviet position on monitoring, US analysts noted that the primary Soviet concern was military vulnerability and the additional targeting information that monitoring would
reveal. 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 in 1984, the Defense Department supported a proposal for extensive inspections of military or government facilities, while intelligence community, Joint Chiefs, and State Department opposed arrangements that would give the other states access to sensitive information beyond that relevant to chemical weapons.

During most of the US-Soviet arms race, closed deals were possible only when unilateral monitoring was sufficiently effective. The online appendix details a case comparison: the 1964 Freeze negotiations failed while the subsequent, similar Strategic Arms Limitation Talks (SALT I) succeeded because unilateral monitoring—in the form of satellite surveillance—improved enough in the interim to make a closed deal viable. On the other hand, the 1987 INF Treaty, which committed the superpowers to eliminate all land-based intermediate- and shorter-range missiles, was an unusual instance of an open deal. The treaty was concluded when the superpowers dampened severity of the tradeoff by designing inspection provisions that gave up some transparency in order to preserve safety.

INF 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. Inspections of Soviet missile production facilities might reveal the technical details or numbers of missiles not slated for elimination. US stealth technology, radar, and industrial processes would have been of interest for Soviet intelligence. 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).

To have an agreement, inspections needed to create just enough transparency to assure compliance, but not so much that safety would be threatened. In early negotiations, the US had pushed for “anytime anywhere inspections” which would allow nearly unfettered access to any missile-related facility. But by the end, the US backed away from these demands and supported limited inspections which would allow access only to declared facilities, and no access to production facilities. This stance was adopted with the understanding that the USSR would be able to hide some violations if it chose to do so. At one point, the USSR even supported a more intrusive option than the US would accept. In his memoir, chief INF negotiator Maynard Glitman explicitly referred to the verification debate as a tradeoff between observing Soviet behavior and threats to US non-treaty interests. The right balance for inspections access was a critical decision without which the treaty could not proceed, and it was a sticking point until the very last stages of negotiation. The ultimate agreement saw both sides backing away from the most extreme options—anytime anywhere inspections and the inclusion of all production facilities—and settling on an inspections regime which, while still highly intrusive, protected the vulnerability of key military installations.

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. The transparency-safety tradeoff meant that options for closed deals were limited and open deals could only be attained with significant compromises that were difficult to negotiate.

**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 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. 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. 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.\footnote{The tradeoff is not completely absent even in the monitoring arrangements that were made. Long (April 3, 2015) argues that these will provide valuable targeting intelligence to the US.}
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.

References

Baliga, Sandeep and Tomas Sjöström. 2008. “Strategic Ambiguity and Arms Proliferation.” 


British Cabinet. 1932. “Memorandum by the Admiralty on Subjects other than Direct Naval Limitation.” The National Archives of the United Kingdom, CAB 27/509.
British Cabinet. 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 United Kingdom, CAB 27/505.


