American Political Science Review (2020) 114, 2, 342–355
doi:10.1017/S000305541900073X

Why Arms Control Is So Rare

ANDREW J. COE *Vanderbilt University*
JANE VAYNMAN *Temple University*

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 security. We present evidence that this trade-off has undermined arms control in three diverse contexts: Iraq's weapons programs after the Gulf War, great power competition in arms in the interwar period, and superpower military rivalry during the Cold War. These arms races account for almost 40% of all global arming in the past two centuries.

Why is arms control so rare? Today and in the past, states arm largely free of agreed constraints. This results in a tremendous loss of wealth: about 2.2% of the world's gross domestic production—$2.7 trillion—was devoted to its militaries in 2016.[1] That 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, its control has received relatively little attention from scholars. The absence of arms control is mysterious for exactly the reason war is. Like war, arming is a costly measure states use to wage their disputes. Arming improves a state's prospects in war and thereby strengthens the state's coercive power, enabling it to better secure its interests. However, just like with war, negotiation could be employed to reach the same outcome as would result from arming but without arming's cost.[3] Why do states arm when there are agreements to limit arms that would leave them strictly better off? In fact, arming is even more puzzling than war, because unlike war, arming is not rare. To the contrary, arming is ubiquitous around the world and throughout history. To explain why arming is so common, we must explain why deals to limit it are so rare.

An obvious answer to why states rarely agree to arms control is that they fear an opponent will cheat in secret and then exploit the resulting advantage. Recent theories of arming assume that states have conflicts of interest and that they cannot perfectly observe each other's arming, which encourages them to arm covertly and thereby undermines arms control (Baliga and Sjöström 2008; Bas and Coe 2016, 2018; Benson and Wen 2011; Debs and Monteiro 2014; Fearon 2011, 2018; Jackson and Morelli 2009; Meirowitz and Sartori 2008; Powell 1993).[4] Earlier work on the security dilemma argued that even if states lacked conflicts of interest, arming could arise out of one side's fear that the other had revisionist interests and so would cheat on any agreed limits.[5]

This fear of covert cheating certainly accords with empirical experience, in which states routinely conceal the details and even overall extent of their arming. However, it cannot actually explain why arms control does not occur. Unlike the balance of power or resolve, arming is not inherently hard to observe. One side can credibly reveal its compliance with arms control by simply allowing the other side sufficient access to its territory. Thus, most existing theories feature arming only because they do not consider the possibility of deals to limit it that utilize monitoring to make cheating observable. With sufficient monitoring, cheating on an agreement will be immediately detected and responded to, so that there is no need for trust and it does not matter whether either side has revisionist interests. Agreements with sufficient monitoring should thus provide an escape from the security dilemma. Why are they nonetheless often not made?

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 security

Andrew J. Coe, Assistant Professor, Department of Political Science, Vanderbilt University, andrew.coe@vanderbilt.edu.
Jane Vaynman, Assistant Professor, Department of Political Science, Temple University, jane.vaynman@temple.edu.

We thank Taylor Dalton for assistance with this research and James Fearon for comments on an earlier draft.

Received: April 23, 2018; revised: June 15, 2019; accepted: November 5, 2019; First published online: December 18, 2019.

[1] This is the most recent figure from the CIA World Factbook.
[2] Of course, not all military spending is lost to the civilian economy. Even if only half is, the remainder would still suffice to pay for all these things, as we document in the Online Appendix.
[3] See Fearon (1995) for the seminal explication of this inefficiency puzzle of war.
[4] Fearon (2011, 2018) and 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.
[5] See Kydd (1997, 2000) for reviews and formalizations of this earlier literature.

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 secure 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 trade-off between transparency and security. If this trade-off is severe, it will be impossible to satisfy both requirements at once. Any deal that is transparent enough to assure that one side complies with the deal may also shift the balance of power so much that the other side reneges 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. When this is true, no arms control deal will be viable.

Empirically, the transparency–security trade-off has severely undermined international attempts at arms control. We show first that the international effort to restrain Iraq from pursuing weapons of mass destruction (WMD) foundered on this trade-off. 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 United States 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 trade-off bedeviled attempts to stop conventional arms competition among the great powers between the world wars. 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 trade-off 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 trade-off shaped Cold War negotiations between the superpowers over limiting their arms race. Primary sources indicate that this trade-off played a role in virtually every negotiation, from the very beginning of the Cold War through to its end. Most agreements were limited to arms that could be monitored unilaterally, precisely because the trade-off inherent in inspections was severe. However, in the rare situations where the trade-off was assessed as relatively mild, deals with inspections were made.

The most recent generation of studies on arms control does not consider the transparency–security trade-off. 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 (2014) examines why only some agreements have verification provisions, but relies on a more general conception of the costs of transparency. Our trade-off generates a new "disclosure dilemma" that is analytically similar to those identified by Carnegie and Carson (2018, 2019), 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.[6]

An earlier generation of work on arms control ignored or underplayed the importance of the transparency–security trade-off. One branch of this literature investigated the optimal design of verification but took for granted that monitoring would not be certain to catch cheating and so ignored our trade-off (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 impediment to arms control (Dunn 1990; Gallagher 1997, 1999; Goldblat 2002; Krass 1985). More recently, Cameron (2018) argues that beliefs about nuclear superiority influenced the willingness of the United States to pursue arms control during the Cold War, and Maurer (2018) emphasizes that the United States favored limits on types of arming in which the USSR was advantaged. This research uncovers important aspects of states' attitudes toward arms control, but it does not suffice to explain why agreements were rarely made. The domestic politics of arms control is vexed only because the transparency–security trade-off is typically severe and so leads each side to refuse monitoring intrusive enough to satisfy the other side's hawks that cheating will be deterred. Moreover, even believers in nuclear superiority should still prefer that it be achieved at the lowest possible cost, so that such beliefs alone are not enough to prevent deals being made.

MODEL SETUP

We proceed to build a model of the interaction between a state that might arm itself and another that might monitor it under a deal to limit arming. We show in the Online Appendix that similar results hold when both sides can arm and be monitored. A (the monitoring state, referred to as feminine) and B (the arming state, masculine) bargain over revisions to a prior settlement

[6] 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 countermeasures.

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.[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 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. 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.[9] 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 $W_o^A(\tau_c) = W_c^A$ and $W_o^B(\tau_c) = W_c^B$. 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.[10] 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-arming to match B's arming. All 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. 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 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. Being open to monitoring obviously makes arming 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 $W_o^A \geq W_c^A$) and reduce the costs of this outcome ($D_o^w \leq D_c^w$), as it makes possible a more effective reaction by A to B's arming. Inspectors may include spies from the monitoring side who can recruit sources, 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–security trade-off and the key to our results.[11]

Our model takes the level of monitoring to be a unilateral, binary choice by the arming side (either τ_c

[7] 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 innocuous.

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

[9] 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, because compliance might still lead to punishment after a false positive signal.

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

[11] We assume that the information exposed by cooperative monitoring cannot instead advantage the monitored state and disadvantage the other state. Any information that can be revealed and whose disclosure would advantage the state possessing it *would* be revealed, whether there is monitoring under an arms control deal or not. It thus cannot affect whether a deal is viable.

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 that, under certain conditions, a deal would not be viable for *any* degree of openness. We also assume that the implementation of monitoring is not itself costly. Although inspectors and other monitoring apparatus do have a budgetary cost, this is typically negligible compared with the costs of arming and conflict.[12] Finally, for simplicity we take the choice of arming to be binary—some or none—but we expect similar results could be obtained with a continuous choice of arming level.

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–security trade-off.[13] 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, no arming or conflict would occur, and there would be no need for any deal. We discard this uninteresting case. Our first result shows that any no-deal equilibrium will be costly for the two sides.

Proposition 1. *In the absence of a deal, B will remain closed to monitoring and invest given the chance, and A may respond with counter-arming or war.*[14]

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. 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 conflict than was required before he armed. This leads B to invest, so that A has cause to worry that his effort will eventually succeed, after which A will have to offer more. 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. 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.

[12] Theoretically, a high enough budgetary cost for implementing monitoring could render a deal too costly to be viable, even without a transparency–security trade-off. Although such a high cost is possible, we suspect it is rare empirically. Incorporating a budgetary cost below this threshold would not qualitatively alter our subsequent results.

[13] 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.

[14] For brevity, the precise conditions governing A's response are relegated to the Online Appendix.

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 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* $\delta[\lambda + \tau_c(1-\lambda)]S \geq -k + \delta\lambda\left[W_n^B - V_p^B\right]$.

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 $\left(\delta\lambda\left[W_n^B - V_p^B\right]\right)$ 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 does not 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 leaves 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 will occur if and only if unilateral monitoring is good enough relative to the temptation to cheat.*

The higher the level of monitoring and the lower the temptation to cheat, the easier it is to deter B from cheating, because cheating is less beneficial if successful and more likely to incur punishment if unsuccessful. This reduces the reward A must offer to ensure compliance and makes a deal more likely 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 relative to the temptation to cheat (labeled $\underline{\tau}$) 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."

This result has clear empirical implications. The more quickly and reliably that B can arm (equivalent to higher λ) and the larger the impact successful arming will have (higher $W_n^B - V_p^B$), the more closely B must be monitored to assure his compliance, and the less likely is a closed deal. For example, the more proficient a state is in nuclear technology, the higher will be the transparency requirement for a nonproliferation deal, because a covert effort to develop nuclear weapons is more likely to succeed quickly. An agreement to instead limit a proficient state to a hundred nuclear weapons need not be monitored as closely as one to limit it to zero, because the impact of cheating on the balance of power will be larger in the latter case. Moreover, if the capabilities being limited have close non-military analogues (i.e., are "dual-use"), or if A's intelligence assets are more limited, then it will be more difficult to distinguish compliance from cheating (equivalent to lower τ_c), and the transparency requirement will again be harder to meet. Warships should be easier to limit in a closed deal than similarly impactful cyber weapons because only the former are large physical objects that are easily distinguishable from civilian analogues.

If A's unilateral monitoring of B's arming is not good enough to meet the transparency requirement ($\tau_c < \underline{\tau}$), then B cannot commit to refrain from investing in arms. A will therefore 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 or providing relevant information. As a result, A would be better able to detect an investment in arming. 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* $\delta[\lambda + \tau_o(1-\lambda)]S \geq -k + \delta\lambda\left[W_n^B - V_p^B\right]$ *and* $1 + \delta W_o^A - k + \delta\left[\lambda W_n^B + (1-\lambda)V_c^B\right] \leq \frac{1}{1-\delta}$.

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 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 to monitoring. It is identical to the condition for a closed deal in Proposition 2, with two exceptions. First, τ_c is replaced by τ_o to reflect that B is open here. Second, the values of punishment for the two sides (V_p^A and V_p^B) are modified to reflect that A might punish B for detected cheating by resorting to costly conflict while B is open to monitoring (giving W_o^A and W_o^B). Because B's openness to monitoring may have given A a military advantage (if $W_o^A > W_c^A$), A might prefer this punishment to the one he would impose under a closed deal (yielding V_c^A and V_c^B, as in Proposition 2).

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. This would cost k now but yield W_n^B next period if the investment succeeded (with probability λ) and V_c^B if it failed (with probability $1-\lambda$). For B, the downside of this option is that, because it is overt, A's punishment will be immediate—if B instead cheated covertly, it might go undetected for a while, so that B's arming is more likely to succeed. The upside is that it prevents A from resorting to costly conflict while B is open to monitoring. This is advantageous to B only if costly conflict while open is a more tempting option for A to punish B or renege on the deal herself (i.e., $W_o^A > V_c^A$), otherwise there is nothing for B to gain from closure. The second inequality in the proposition assures that it is possible for a deal to give each side a value at least as good as these outside options: W_o^A (for A) and $-k + \delta\left[\lambda W_n^B + (1-\lambda)V_c^B\right]$ (for B) should not sum to more than the total value available ($\frac{1}{1-\delta}$).[15]

These two conditions imply a critical but subtle obstacle to the viability of any open deal, which we call the transparency–security trade-off. 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 trade-off 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 $W_c^B = W_o^B$), so that there is no trade-off.

Corollary 2. *If there is no transparency–security trade-off, a deal will always be made. There will be 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 needed.

[15] The "1" in the second inequality comes from the fact that B can only cheat after having accepted A's offer, so that the value of the game for that period alone is 1. The δ attached to W_o^A reflects that A can only resort to costly conflict in the round after B has opened.

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–security trade-off, B is happy to do so because the deal leaves him strictly better off. The key here is that increasing transparency by opening to monitoring has no downside for B.

This result carries a remarkably counterintuitive implication. If there is no trade-off, then a deal can be made *no matter how great B's temptation to cheat is.* Even if B could swiftly acquire a decisive new military capability if he chose to arm, and even if A's response to that arming would be slow and ineffective, so that a serious shift in power would occur, a deal will still be viable. To be sure, a situation where B can arm speedily but A can only respond slowly is advantageous for B and disadvantageous for A. However, if monitoring under a deal can be made perfect, then the expected outcome for both sides of B cheating on a deal will be the same as the expected outcome in the deal's absence: B will arm quickly, A will instantly know this is occurring, and A will respond exactly as slowly. Making a deal thus would not leave A any worse off militarily, or B any better off militarily, than each would be in the deal's absence. But making a deal would save both sides the costs of B's arming and A's response to it, and as long as monitoring itself does not alter the military balance, a deal with perfect monitoring will be strictly preferred by both sides to going without a deal.

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–security trade-off and a mutually acceptable deal is less likely to exist.

Corollary 3. *Suppose unilateral monitoring is not good enough relative to the temptation to cheat to support a closed deal. An open deal will occur if and only if the transparency–security trade-off is not too severe ($\underline{\tau} \leq \overline{\tau_o}$, where $\overline{\tau_o}$ is the largest value of τ_o such that $1 + \delta W_o^A(\tau_o) - k + \delta\left[\lambda W_n^B + (1-\lambda)V_c^B\right] \leq \frac{1}{1-\delta}$).*

With a transparency–security trade-off, 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–security trade-off 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 secure enough for B. This is the "security requirement": τ_o can be no more than $\overline{\tau_o}$, the highest level of openness that would still not lead to A reneging on the deal.[16]

Obviously, these two requirements are in tension. If the trade-off between them is mild, so that security does not decrease much as transparency increases (i.e., $W_o^A(\tau_o)$ does not rise quickly in τ_o), then there will be a range of levels of openness ($[\underline{\tau}, \overline{\tau_o}]$) 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 trade-off is severe, so that security declines rapidly as transparency increases ($W_o^A(\tau_o)$ rises rapidly in τ_o), there is no level of openness that would both assure A of B's compliance and assure B of A's.

The severity of this trade-off should depend on attributes both of the agreed limits and of the state to be monitored. The physical nature of the capability to be limited, and the specific restrictions agreed, determine what is entailed in reaching a certain level of monitoring τ_o of compliance with those limits. How many places have to be inspected, how often, how closely? A capability that is smaller, can be produced or deployed more quickly and at more sites, or more closely resembles civilian or other non-restricted capabilities, will require more intrusive inspections to reach a given level of confidence in detecting a violation.

Characteristics of the state being monitored can determine the value of the information that might be inadvertently revealed by a given level of monitoring. What other secrets does the state have that are "colocated" with the limited capability and might be exposed by monitoring? Monitoring within a more generally closed state, or of a capability that is produced or deployed at government facilities, especially weapons factories or military bases or facilities involved in internal regime security, will entail higher security risks at any given level of monitoring. By contrast, if few secrets are colocated with the capability to be limited, then increasing transparency poses less threat to security and the severity of the trade-off decreases.

To illustrate the determinants of the trade-off's severity, consider an agreement to ban the production of fissile material versus a ban on weaponizing fissile material. The former might involve inspections of nuclear reactors, reprocessing plants, and enrichment facilities. These are typically large and few in number. They produce fissile material slowly, and this production is difficult to disguise. By contrast, weaponizing fissile material that is freely produced could require inspections at many military bases where weaponization could be undertaken. It is a relatively small activity that can be done quickly and disguised easily as conventional explosives testing. Hence, the latter ban will require more intrusive monitoring to reach a given probability of detecting cheating, and so will face a more severe transparency–security trade-off.

Now consider two different states that might agree to an arms control deal with cooperative monitoring. In one state, a free press reports on economic, military, and political affairs and the relevant sites for inspections are owned and operated by civilian firms that play no role in external or internal security. Inspections would pose few security risks for this state because the sites are not colocated with other secrets and because much information about the state is already publicly known. In the other state, the government controls the press and conceals many aspects of the state's affairs, the relevant

[16] The definition of $\overline{\tau_o}$ in the corollary is taken directly from the second inequality in Proposition 3.

facilities are owned and operated by the military, and the military plays important roles in both internal and external security. Inspections will entail greater security risks for this state because they threaten to reveal information that the government would prefer to hide from foreign or domestic enemies. These secrets might include the quantity and quality of the state's other military capabilities, the condition of its economy, the extent of corruption in government, and the presence of dissension within the military or bureaucracy. The second state will therefore confront a more severe transparency–security trade-off than the first.

The transparency–security trade-off renders a deal under open monitoring harder to make 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 or arming.

This result builds upon the conclusions of prior theories of arming. Powell (1993) shows that the costs of arming and the shifts in power that attend it 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 asymmetric information in itself 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; and the number of overflights allowed. The intensity of the transparency–security trade-off 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 relative to the temptation to cheat to convince the monitoring side that the arming side's compliance can be assured, there will be a closed deal to limit this arming (Corollary 1). If instead the monitoring side believes that cooperation is required to assure sufficient transparency, and this level is also seen as secure enough by the arming side, there will be an open deal (Corollary 2). Finally, if the cooperation necessary to assure transparency to one side is seen as too threatening to the security of the other, there will be no deal at all (Corollary 3). In sum, the transparency–security trade-off 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 trade-off between transparency and security is a serious and common impediment to making arms control deals empirically. We examine the negotiations between the United States and Iraq over the latter's pursuit of weapons of mass destruction; the efforts by the great powers to constrain their conventional arms competition between the World Wars; and finally, some of the superpowers' Cold War attempts to limit their arms race. We chose these settings because they vary widely in time period, arms to be limited, and characteristics of the involved states, allowing us to assess how broadly the theory applies. Collectively, our cases allow us to test all three of the theory's observable implications about whether a deal is made, what kind of monitoring it features, and which arms it limits.

In each case, we investigate policy makers' perceptions of the temptation to cheat on a deal relative to their ability to unilaterally monitor compliance, as well as of the presence and severity of the trade-off. In the U.S.–Iraq case, we show that unilateral monitoring was perceived to be inadequate and the trade-off was perceived to be severe, and that this prevented a deal being made. The interwar setting shows how the trade-off can determine which arms are limited: deals on naval but not air or ground forces were made, because only the former were easy to monitor unilaterally, and cooperative monitoring of the latter posed a severe trade-off. Finally, we study two pairs of cases from the Cold War. The two rounds of negotiations toward the INF Treaty demonstrate the effect of changing the perceived trade-off from severe to mild: an open deal becomes viable. The Freeze and SALT I negotiations illustrate the effect of increasing the perceived ability to unilaterally monitor compliance: a closed deal becomes viable.

Because we find that the theory helps to explain arms control outcomes across a variety of settings, we conclude that it is broadly applicable. Because the arms races occurring in these settings account for 40% of the world's military spending for the last two centuries, the fact that a severe trade-off sometimes arose in all of them suggests that the trade-off is common enough to be responsible for a substantial portion of all failures to control arms.[17]

[17] Figures are computed from the National Military Capabilities dataset and are adjusted for inflation.

Iraq and the United States, 1990–2003

The United States imposed sanctions on, encouraged rebellion within, undertook limited strikes against, and eventually invaded Iraq to overthrew 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 United States and Iraq. Both would have been better off had they instead struck a deal in which Iraq disarmed and enabled the United States to verify this, and in exchange the United States did not sanction, strike, or invade Iraq. Why wasn't a deal made?

We will first argue that Iraq's temptation to cheat was very high relative to the United States ability to unilaterally monitor its compliance, so that as Corollary 1 predicts, no closed deal was made. We then demonstrate that there was a severe transparency–security trade-off, which Corollary 3 predicts should rule out any open deal. We draw on the arguments and evidence assembled in Koblentz (2018) that inspections posed a very serious threat to Iraq's regime, combining these with our own arguments about unilateral monitoring and Iraq's temptation to cheat to explain why no deal was viable.[18]

Iraq's temptation to cheat on any deal was high because acquiring the prohibited arms would have had a large impact on the outcome of conflict, and an effort to develop them would likely have succeeded quickly. Both Saddam and the United States anticipated that Iraq's acquisition of WMD—especially but not only nuclear weapons—would radically shift the balance of power in Iraq's favor (Brands and Palkki 2011; Iraq Survey Group 2004, Key Findings, 24–8, 33; Pollack 2002, 175–7, 249–53, 268–70, 272–6). Moreover, Iraq had possessed biological and chemical weapons and had come very close to acquiring nuclear weapons before the Gulf War (Richelson 2007, 464). Although these programs were dismantled after the war, Iraq maintained the human capital needed to reconstitute them (Iraq Survey Group 2004, Key Findings, 24, 44, 49, 51). Iraq's cheating under a deal was therefore likely to result in its acquisition of highly consequential new arms relatively quickly.

The ability of the United States to unilaterally monitor Iraq's nuclear ambitions was inadequate to deter this temptation to cheat. Iraq had successfully hidden the extent and progress of its nuclear program from the United States 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, 45–7). Thus, the United States 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 U.S. 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, 46–7, 49, 81–2, 114–5, 154–6, 169–70). The establishment of UNSCOM was motivated by the recognition that Iraq's disarmament could not be assured otherwise. Consistent with Corollary 1, no closed deal was viable.

Thus, highly intrusive inspections were necessary to provide the transparency required for a viable deal. Debs and Monteiro (2014, 25) argue 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 2018; Harvey 2011). However, there was a sufficiently intrusive set of inspections that, if permanently implemented, would have satisfied the United States 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 (Stiglitz and Bilmes 2008). 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 United States was therefore certainly *possible* with cooperative monitoring.

Unfortunately, achieving such high transparency would seriously endanger Iraq's regime and the survival of Saddam personally, so that the transparency–security trade-off was severe. During this period, the United States engaged in a covert campaign to help Saddam's domestic opposition to overthrow his regime (Koblentz 2018, 23–4).[19] Though the resulting popular uprisings, coup plots, and assassination attempts against Saddam failed (8–9), they nonetheless demonstrated a clear and dire threat to the regime (15–7). Crucially, the inspections necessary to verify Iraq's abstention from WMD also increased the chance that this U.S. campaign would succeed and the regime would be overthrown, because information inadvertently revealed by those inspections could be used to assist the campaign. The CIA penetrated prior inspection teams and reportedly used inspection visits to coordinate with coup plotters and opposition elements (24–5). 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

[18] The conventional wisdom for why Iraq did not fully reveal its lack of WMD is that it sought to maintain ambiguity to deter Iran. Koblentz (2018) shows there is no evidence that corroborates this theory and ample evidence that contradicts it.

[19] We henceforth use only page numbers in citing this article.

about Saddam's security measures and how to defeat them (24–5), which the United States could then pass on to Saddam's domestic opponents. Iraq's Special Security Office (SSO) 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 (16–7), so that monitoring it would also have yielded intelligence to the United States on the regime's vulnerabilities. Increasing transparency would therefore come at a very high cost to the Iraqi regime's security.

Saddam was keenly aware of the trade-off between transparency and security involved in the inspections, and his behavior was driven by it (3–4, 22–6). 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 (22–3). As a result, the most contentious issues in negotiations between Iraq and the international community over inspections were those for which the trade-off was most severe, especially inspections of Saddam's palaces and of the SSO (27–8). Almost all the instances in which an inspection visit was obstructed involved such sites (9–12, 21–2).

The United States was largely unaware of how pressing this trade-off 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 security (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 that Saddam sought to threaten with WMD capabilities (2, 5–12).

Given Iraq's high temptation to cheat, the inability of the United States to unilaterally monitor compliance, and the severe transparency–security trade-off, Corollary 3 stipulates that no deal would be mutually acceptable. Indeed, during this period, the United States constantly demanded more transparency, Iraq resisted to protect the regime's security (4), and no deal was made. Counterfactually, were it not for the trade-off, Iraq would have had no reason to refuse even the most intrusive inspections. The United States would therefore have been able to assure itself of Iraq's compliance, and the war would have been avoided.

The Interwar Arms Negotiations

Between the two World Wars, the great powers attempted to limit their conventional arms competition. 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-1930s.[20] Although limits on air and ground forces were discussed in international conferences, no agreements were made, and competition in these arms continued.

Our theory helps to explain why agreements were made on naval but not other forces. We will argue that the temptation to cheat on limits on large warships was low relative to the great powers' ability to unilaterally monitor compliance with those limits, so that as Corollary 1 predicts, a closed deal was made. By contrast, the temptation to cheat on limits on air and ground forces was high relative to the ability to monitor those limits unilaterally, so that a closed deal could not be made. Because inspections to render compliance more observable would have exposed serious vulnerabilities, the transparency–security trade-off was severe, so that Corollary 3 predicts no open deal to limit these forces would be viable either.

First consider naval forces. Because these were slow to build and any substantial cheating was easily detected, the temptation to cheat on agreements to limit them was low relative to unilateral monitoring ability. Limits on capital ships could be monitored quite easily through unilateral means such as espionage and media reports (Goldman 1994, 178–80). These ships were massive and typically built in the open, so that spies could readily determine whether construction was ongoing (Overy 2016, 129–30). Moreover, they required years to complete, leaving plenty of time for detection (Maiolo 2016, 62) and therefore also for states to react effectively to cheating. The British Admiralty assessed that "with regard to numbers of ships or the size of guns carried, investigation [i.e., inspection] 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" (Flint 1921, 930). The United States also "regarded the capital ship as easy to count and difficult to hide" (Kaufman 1990, 57). 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 1921–1922, 148). Espionage could not detect small tonnage violations, but these were seen as marginal because "large excess would be readily detected" without inspection (British Cabinet 1932, 4).

Unilateral monitoring therefore sufficed to meet the transparency requirement, and as Corollary 1 predicts, closed deals were made. The Washington and London Naval Treaties limited large warships with no provision for cooperative monitoring and consequently avoided any transparency–security trade-off.

The great powers also considered limiting air and land forces. The temptation to cheat on these limits was high relative to unilateral monitoring ability because the underlying capabilities were dual use: those in civilian use could be quickly converted to military use, and the two were difficult to distinguish. For land forces, conscription provided for rapid expansion (British Cabinet 1921–1922, 140–6). 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

[20] Subsequent political changes in Germany and Japan led them to abandon these treaties and made further limitations untenable (Maurer 1994, 289).

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 United States "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" (Fuller 1938, 314).

Limits on air forces posed similar difficulties because the factories that produced commercial vehicles could be quickly turned to military production. The British noted "the narrow line between commercial and military aviation" (Foreign Office 1921, 475–6) 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, 129–30). The United States "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).

Because the temptation to cheat on limits on air and ground forces was high relative to unilateral monitoring ability, Corollary 1 predicts that no closed deal could be made, and indeed none was. However, it was possible to verify compliance with limits on land and air forces with cooperative monitoring. Indeed, this had been 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" (Cubitt 1921, 927–8). 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 (Foreign Office 1925, 190–7). 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 1921–1922, 146–7), and viewed such inspections as required for any deal on air or ground forces (Kitching 2003, 156, 162). The British, despite being inclined to leniency, nevertheless believed such inspections remained necessary to monitor limits on German land forces (Cubitt 1921, 927–8).

The problem with such inspections was that they would expose serious military vulnerabilities, so that the transparency–security trade-off was severe. Britain opposed inspections out of fear that these would reveal a lack of defense preparations and other secrets (Kitching 2003, 156, 160–1; 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 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" (British Cabinet 1933, 9). For its part, the United States "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).

Consistent with Corollary 3, no open deal was made.[21] British resistance to cooperative monitoring scuttled the various attempts at general disarmament (British Cabinet 1921, 159–61; British Cabinet 1932, 5–6; Kitching 2003, 160–72). 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. Our theory explains why more comprehensive limits were not agreed.

[21] The major powers could instead have agreed to a ban on the manufacture of mechanized vehicles, whether civilian or military. This would reduce the speed with which mechanized forces could be created, lowering the transparency requirement. Violations would be also be easier to detect, because there would be no need to distinguish between military and civilian vehicles. Compliance might then be assured by unilateral monitoring, so that the transparency–security trade-off could be avoided. However, this deal would require sacrificing the value of mechanized vehicles to the civilian economy. If this value is too high, this deal will not be viable. The possible use of costly restrictions on civilian activity to make limits on military capabilities easier to monitor is not captured in our model. Further research is needed on such deals.

Verification during the Cold War

The United States and USSR undertook the most expensive and most dangerous arms race in human history. In their attempts to constrain this race, both sides recognized the transparency–security trade-off, and it was crucial to their evaluation of mooted limits and monitoring. Agreements emerged only in the few instances in which unilateral monitoring sufficed to assure compliance or the trade-off in cooperative monitoring was assessed to be mild. Here, we first show that the trade-off played a role in early considerations of comprehensive arms limits and then summarize our findings across several later, more narrow negotiations during the Cold War. For brevity's sake, documentation of these claims based on secondary and archival sources is relegated to the Online Appendix.

From the very beginning of the Cold War, both sides saw cooperative monitoring as generally necessary to meet the transparency requirement for limits on nuclear and conventional forces, but also perceived serious security risks in such monitoring. A 1951 CIA report on disarmament proposals notes: "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). The report analyzes which side would gain more military advantage through "collateral" information gleaned by inspections. It recommends limiting inspections to protect U.S. information even though inspections are the only way to assure Soviet compliance. Internal Soviet discussions are less readily available, but occasional public statements also reveal an early appreciation of the transparency–security trade-off. In 1959, a Soviet proposal on general disarmament at the UN noted that inspections, though needed, would also reveal information detrimental to security.

Closed deals were made only under the conditions of Corollary 1: when the temptation to cheat was low relative to unilateral monitoring, so that unilateral monitoring sufficed to assure compliance and the trade-off was therefore not operative. In the Online Appendix, we argue that the 1964 Freeze negotiations failed, whereas the subsequent, similar Strategic Arms Limitation Talks (SALT I) succeeded because the development of satellite surveillance improved unilateral monitoring enough to make a closed deal viable.

The negotiations over the Intermediate Nuclear Forces (INF) Treaty, which eliminated all land-based intermediate-range missiles, allow us to test Corollary 2. This predicts an open deal when the temptation to cheat is high relative to unilateral monitoring ability, but the transparency–security trade-off is mild. The first round of negotiations was stymied by the two sides' perception that the transparency-security trade-off was severe. However, during the second round, altered inspection procedures were devised to achieve the required level of transparency while minimizing risks to security, so that the trade-off was reassessed as mild. Just as Corollary 2 predicts, this led to the INF Treaty, an open deal.

Unilateral monitoring was judged insufficient to overcome the temptation to cheat on the INF limits. Earlier agreements covered missiles housed in fixed silos, and so could be reliably monitored unilaterally via satellite. By contrast, INF missiles were smaller and carried on mobile launchers, which could move around easily and at night, escaping satellite detection. Inspections were thus needed to ensure that existing missiles were eliminated and no new ones were produced. In the first round of negotiations (1980–83), the United States proposed "anytime anywhere inspections," which would allow nearly unfettered access to any missile-related facility to achieve transparency.

However, these same inspections also risked exposing important military secrets, so that the transparency–security trade-off was perceived to be severe. Inspections of Soviet missile production facilities might reveal the technical details and quantity of missiles not prohibited under the agreement. On the U.S. side, stealth technology, radar, and industrial processes would be subject to Soviet espionage. The USSR refused early inspections proposals, believing that they would be used to "discover the strengths, weaknesses, and vital characteristics of Soviet weapons." Tellingly, the U.S. policy makers ultimately came to the conclusion that they could not accept their own proposal because of the risks to security. Nongovernmental experts traditionally supportive of arms control raised concerns that highly intrusive inspections would mean "Soviets crawling around our most sensitive production facilities" and leave the United States at a relative disadvantage. The CIA even recommended doing away with inspections.

During the second round of negotiations (1985–87), a new approach to inspections was devised and both sides reassessed the severity of the trade-off. In 1986, the United States began to focus on the prospects for more limited inspections and subsequently proposed intrusive inspections but with access to only one production facility. The United States also devised a way to monitor this facility with a sensor that would reveal whether an exiting missile was of the banned type, but not reveal the technical details of non-banned missiles produced at the same location. Although it was understood that the USSR would be able to hide some violations under this inspections proposal, the required level of transparency could be still be achieved, and with much lower security risks than the earlier proposal of "anytime anywhere" inspections. Thus, the United States came to view the transparency–security trade-off as mild.

The USSR also overcame its long-standing opposition to inspections, with serious proposals for on-site inspections emerging during this period. This change derived largely from revised assessments of the risks that inspections would pose to Soviet military capabilities. As documented in memoirs and Politburo accounts, the political leadership concluded that the military's resistance to inspections had more to do with obfuscating bureaucratic excess and incompetence than with protecting Soviet military power vis-á-vis the United States. Inspections of the kind proposed by INF

negotiators might embarrass the military but would have little effect on the balance of power. Thus, the USSR now perceived the transparency–security trade-off to be mild.

A mutual understanding that the trade-off was mild was critical to making the deal. Debates over this issue continued through the very last stages of negotiation, delaying the treaty even though other terms—such as the level and scope of limitations—were already settled. Chief U.S. negotiator Maynard Glitman explicitly referred in his memoir to the verification debate as a trade-off between observing Soviet behavior and tolerating threats to U.S. military secrets. The ultimate INF Treaty eschewed anytime-anywhere inspections because they achieved transparency but created unacceptable risks to security. Instead, this open deal featured inspections, which, although still highly intrusive, protected the secrets of key military installations and were therefore safer, in accordance with Corollary 2.

Both the United States 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 an exorbitantly costly arms race. The transparency–security trade-off prevented many attempts at arms control from ever reaching agreement.

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 trade-off involved in monitoring: transparency is required to assure one side of the other's compliance with arms limits, but transparency might also reveal vulnerabilities that could be exploited by the first side in an arms race or war. If this transparency–security trade-off is severe, then any cooperative monitoring that would assure one side would be unsafe for the other. If unilateral monitoring alone 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 trade-off 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–security trade-off 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 trade-off altogether, because both transparency and security can be achieved. In practice, such a third-party can only be effective if both sides trust that it is able and willing to uphold its duty. A state may fear that the third-party is not impartial, so that it could not be counted on to reveal discovered evidence of cheating. 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 United States 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 over time, with most occurring in the latter half of the twentieth century. We conjecture that this increase is related to the rise of electronic communication and overhead imagery, which provided new avenues for unilateral monitoring of arms control deals. Spy satellites and the unilateral monitoring they enabled were key to 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 United States believed its general openness left it at a disadvantage to the more closed USSR, leading the United States to favor cooperative monitoring and the USSR to resist it. Generalizing, more open dyads should be more likely to make arms control agreements than less open dyads, conditional on the members of the dyad being adversaries.

A better understanding of the transparency–security trade-off might also inform assessments of new arms control deals. Consider the recent deal between the United States 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 those military bases might enable the United States not only to learn whether Iran was complying with the deal but also to gain valuable targeting information that could be exploited in a future conflict (Long 2015). 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–security trade-off, so that improvements would be acceptable to both sides.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S000305541900073X.

REFERENCES

Baliga, Sandeep, and Tomas Sjöström. 2008. "Strategic Ambiguity and Arms Proliferation." *Journal of Political Economy* 116 (6): 1023–57.

Bas, Muhammet A., and Andrew J. Coe. 2016. "A Dynamic Theory of Nuclear Proliferation and Preventive War." *International Organization* 70 (4): 655–85.

Bas, Muhammet A., and Andrew J. Coe. 2018. "Give Peace a (Second) Chance: A Theory of Nonproliferation Deals." *International Studies Quarterly* 62 (3): 606–17.

Benson, Brett, and Quan Wen. 2011. "A Bargaining Model of Nuclear Weapons Development and Disarmament." In *Causes and Consequences of Nuclear Proliferation*, eds. Robert Rauchhaus, Matthew Kroenig, and Erik Gartzke. New York, NY: Taylor and Francis, 45–62.

Brands, Hal, and David Palkki. 2011. "Saddam, Israel, and the Bomb: Nuclear Alarmism Justified?" *International Security* 36 (1): 133–66.

British Cabinet. 1921. Washington Disarmament Conference. Memoranda. Vol. I. The National Archives of the United Kingdom, CAB 30/5.

British Cabinet. 1921–22. Conference on Limitation of Armaments, Washington, 1921–22: Plenary Sessions. Stenographic Notes. The National Archives of the United Kingdom, CAB 30/3.

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.

Cameron, James. 2018. *The Double Game: The Demise of America's First Missile Defense System and the Rise of Strategic Arms Limitation*. New York, NY: Oxford University Press.

Carnegie, Allison, and Austin Carson. 2018. "The Spotlight's Harsh Glare: Rethinking Publicity and International Order." *International Organization* 72 (3): 627–57.

Carnegie, Allison, and Austin Carson. 2019. "The Disclosure Dilemma: Nuclear Intelligence and International Organizations." *American Journal of Political Science* 63 (2): 269–85.

Central Intelligence Agency. 1951. Special Estimate: Intelligence Implications of a Census and Verification of Armed Forces and Armaments. FOIA Electronic Reading Room: National Intelligence Council (NIC) Collection. Document Number (FOIA)/ESDN (CREST): 0000170378.

Coe, Andrew J. 2018. "Containing Rogues: A Theory of Asymmetric Arming." *The Journal of Politics* 80 (4): 1197–210.

Coe, Andrew J., and Jane Vaynman. 2015. "Collusion and the Nuclear Nonproliferation Regime." *The Journal of Politics* 77 (4): 983–97.

Cubitt, B. B. 1921. Annex III to No. 815: War Office to Foreign Office. Documents on British Foreign Policy 1919–39 Ser. 1, Vol. 16, Ch. 7.

Dai, Xinyuan. 2002. "Information Systems in Treaty Regimes." *World Politics* 54 (4): 405–36.

Dai, Xinyuan. 2007. *International Institutions and National Policies*. New York, NY: 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.

Downs, George W., and David M. Rocke. 1990. *Tacit Bargaining, Arms Races, and Arms Control*. Ann Arbor, MI: University of Michigan Press.

Dunn, Lewis A. 1990. "Arms Control Verification: Living with Uncertainty." *International Security* 14 (4): 165–75.

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. 2018. "Cooperation, Conflict, and the Costs of Anarchy." *International Organization* 72 (3): 523–59.

Flint, Alexander. 1921. Annex V to No. 815. Documents on British Foreign Policy 1919–39 Ser. 1, Vol. 16, Ch. 7.

Foreign Office. 1921. The Marquess Curzon of Kedleston to Mr. Balfour (Washington Delegation) No. 7 Telegraphic [A 8510/18/45]. Documents on British Foreign Policy 1919–39. Ser. 1, Vol. 14, Ch. 6.

Foreign Office. 1925. Memorandum on the Future of Air Control in Germany. Documents on British Foreign Policy 1919–39 Ser. 1A, Vol. 1, Ch. 2.

Fuller, Joseph V. ed. 1938. *Papers Relating to the Foreign Relations of the United States, 1922, Volume I*. Washington, DC: United States Government Printing Office. Document 88.

Gallagher, Nancy W. 1997. The Politics of Verification: Why 'how Much?' Is Not Enough. *Contemporary Security Policy* 18 (2): 138–70.

Gallagher, Nancy W. 1999. *The Politics of Verification*. Baltimore, MD: Johns Hopkins University Press.

Goldblat, Jozef. 2002. *Arms Control: The New Guide to Negotiations and Agreements*. Thousand Oaks, CA: Sage.

Goldman, Emily O. 1994. *Sunken Treaties: Naval Arms Control between the Wars*. University Park, PA: Pennsylvania State University Press.

Harvey, Frank P. 2011. *Explaining the Iraq War: Counterfactual Theory, Logic and Evidence*. New York, NY: 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. Unclassified Version, 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: Columbia University Press.

Kilgour, D. Marc. 1994. "The Use of Costless Inspection in Enforcement." *Theory and Decision* 36 (3): 207–32.

Kitching, Carolyn J. 2003. *Britain and the Problem of International Disarmament: 1919–1934*. New York, NY: Routledge.

Koblentz, Gregory D. 2018. "Saddam versus the Inspectors: The Impact of Regime Security on the Verification of Iraq's WMD Disarmament." *Journal of Strategic Studies* 41 (3): 372–409.

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): 731–52.

Kreps, Sarah E. 2016. "The Institutional Design of Arms Control Agreements." *Foreign Policy Analysis* 14 (1): 127–47.

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): 228–44.

Lindsey, David. 2015. "Military Strategy, Private Information, and War." *International Studies Quarterly* 59 (4): 629–40.

Long, Austin. 2015. "If You Really Want to Bomb Iran, Take the Deal." *Washington Post* (April 3).

Maiolo, Joseph. 2016. "Between the Two World Wars: Introduction." In *Arms Races in International Politics: From the Nineteenth to the Twenty-First Century*, eds. Thomas G. Mahnken, Joseph Maiolo, and David Stevenson. Oxford, UK: Oxford University Press, 61–8.

Maurer, John D. 2018. "The Purposes of Arms Control." *Texas National Security Review* 2 (1).

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*, eds. Erik Goldstein and John Maurer. New York, NY: Routledge, 267–93.

Meirowitz, Adam, and Anne E. Sartori. 2008. "Strategic Uncertainty as a Cause of War." *Quarterly Journal of Political Science* 3 (4): 327–52.

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, eds. Thomas G. Mahnken, Joseph Maiolo, and David Stevenson. Oxford, UK: Oxford University Press, 115–33.
Pollack, Kenneth M. 2002. *The Threatening Storm: The Case for Invading Iraq*. New York, NY: Random House.
Powell, Robert. 1993. "Guns, Butter, and Anarchy." *American Political Science Review* 87 (1): 115–32.
Richelson, Jeffrey T. 2007. *Spying on the Bomb: American Nuclear Intelligence from Nazi Germany to Iran and North Korea*. New York, NY: W. W. Norton & Company.
Slantchev, Branislav L. 2010. "Feigning Weakness." *International Organization* 64 (3): 357–88.
Stiglitz, Joseph E., and Linda J. Bilmes. 2008. *The Three Trillion Dollar War: The True Cost of the Iraq Conflict*. New York, NY: W. W. Norton & Company.
Vaynman, Jane. 2014. "Enemies in Agreement: Domestic Politics, Uncertainty, and Cooperation between Adversaries." PhD thesis, Harvard University.
Verdier, Daniel. 2008. "Multilateralism, Bilateralism, and Exclusion in the Nuclear Proliferation Regime." *International Organization* 62 (3): 439–76.
Wittman, Donald. 1989. "Arms Control Verification and Other Games Involving Imperfect Detection." *American Political Science Review* 83 (3): 923–45.