"Known Unknowns:

Power Shifts, Uncertainty, and War."

Online Appendix

Alexandre Debs and Nuno P. Monteiro

May 10, 2016*

The Appendix is structured as follows. Section 1 offers proofs of the formal results in the two-period game discussed in the main text. Section 2 offers an expanded presentation of the infinite-horizon game set-up and results. Section 3 offers proofs of the formal results of the infinite-horizon game.

*The first draft of the Appendix, as it was published online, was completed on June 21, 2013. This version, written on May 10, 2016, makes some minor corrections and clarifications, in footnotes 2, 4, 6, 7, and 10.

1 Two-Period Game: Proofs

We say that the effect of militarization is smaller than the cost of investment if

$$\delta\left[w_T(1) - w_T(0)\right] \leq k \tag{1}$$

We say that the effect of militarization is smaller than the cost of preventive war if

$$\delta\left[w_T(1) - w_T(0)\right] \leq 1 - w_T(0) - w_D(0) \tag{2}$$

We say that the signal is sufficiently informative if

$$(1 - p_s)\,\delta\left[w_T(1) - w_T(0)\right] \leq k \tag{3}$$

Proof. (Proof of Proposition 2). Formally, we show[1]

(i) If (1) holds strictly or if (2) holds strictly, then peace prevails.

(ii) If (1) holds, the PBE is: $I_1^* = 0$. *D offers* $z_1^* = w_T(0)$ *unless* $s_1 = 1$ *and (2) fails, in which case D declares war. T accepts any* $z_1 \geq w_T(0)$.

(iii) If (1) fails and (2) holds, the PBE is: $I_1^* = 1$. *D offers* $z_1^* = w_T(0)$ *after any* s_1. *T accepts any* $z_1 \geq w_T(0)$.

To see this, note first that in any equilibrium T accepts any offer $z_1 \geq w_T(0)$.

Consider D's choice between declaring war and offering $z_1 = w_T(0)$. The expected utility for D of a preventive war is $w_D(0) + \delta(1 - w_T(0))$. The expected utility for D of offering $z_1 = w_T(0)$ is at least equal to $1 - w_T(0) + \delta(1 - w_T(1))$. Thus, if (2) holds strictly, D strictly prefers to offer $z_1 = w_T(0)$ and peace prevails.

Consider T's decision to invest in military capabilities. The expected utility for T of investing in military capabilities is at most equal to $-k + w_T(0) + \delta w_T(1)$. The expected utility for T of not investing in military capabilities is equal to $w_T(0) + \delta w_T(0)$. Thus, if (1) holds strictly, T strictly prefers not to invest in military capabilities and as a result D

[1] In non-generic regions of the parameter space, such as (1) and (2) holding with equality, there could be multiple equilibria. In these cases we assume that any player breaks indifference in favor of the efficient action, not investing for T and making a peaceful offer for D.

strictly prefers to offer $z_1 = w_T(0)$, since war is inefficient $(1 - w_T(0) + \delta(1 - w_T(0)) > w_D(0) + \delta(1 - w_T(0)))$. This completes the proof of (i).

Now assume that (1) holds. Then T chooses $I_1^* = 0$. As a result, D prefers to declare war if and only if $w_D(0) + \delta(1 - w_T(0)) > 1 - w_T(0) + \delta(1 - w_T(s_1))$ or if and only if $s_1 = 1$ and (2) fails. This completes the proof of (ii).

Now assume that (1) fails and (2) holds. Since (2) holds, D offers $z_1^* = w_T(0)$. As a result, T chooses $I_1^* = 1$ since (1) fails. This completes the proof of (iii). ■

Proof. (Proof of Proposition 3). Formally, we show:

Assume that (1) and (2) fail.

(i) If (3) holds, the PBE is: $I_1^* = 0$*. D offers* $z_1^* = w_T(0)$ *after* $s_1 = 0$ *and declares war after* $s_1 = 1$*. T accepts any* $z_1 \geq w_T(0)$*.*

(ii) If (3) fails, the PBE is: T invests with probability $q^* = \frac{1}{p_s + (1-p_s)\frac{\delta[w_T(1) - w_T(0)]}{1 - w_T(0) - w_D(0)}}$*. After* $s_1 = 0$*, D offers* $z_1^* = w_T(0)$ *with probability* $r^* = \frac{k}{(1-p_s)\delta[w_T(1) - w_T(0)]}$ *and declares war with probability* $1 - r^*$*. After* $s_1 = 1$*, D declares war. T accepts* $z_1 \geq w_T(0)$*.*

To see this, note that T accepts any $z_1 \geq w_T(0)$. Also, D declares war after $s_1 = 1$ since (2) fails.

After signal $s_1 = 0$, D must choose between the best response to either pure strategy by T. Let D offer $z_1^* = w_T(0)$ with probability r^* and declare war with probability $1 - r^*$ after signal $s_1 = 0$. Let T invest with probability q^*.

Next we show that $q^* < 1$. Indeed, if $q^* = 1$, then D prefers $r^* = 0$, so that T prefers $q^* = 0$, a contradiction. Second, if $q^* = 0$, then D prefers $r^* = 1$, so that T has no incentive to deviate if and only if

$$(1 + \delta) w_T(0) \geq -k + (1 + \delta) w_T(0) + (1 - p_s)\delta(w_T(1) - w_T(0))$$

or (3) holds. If (3) fails, we must have $q^* \in (0, 1)$. This in turn requires that $r^* \in (0, 1)$.

$q^* \in (0,1)$, $r^* \in (0,1)$ are given by the indifference conditions:

$$(1+\delta)\, w_T(0) = -k + (1+\delta)\, w_T(0) + (1-p_s)\, r^* \delta \left(w_T(1) - w_T(0)\right)$$
$$w_D(0) + \delta\left(1 - w_T(0)\right) = (1+\delta)\left(1 - w_T(0)\right) - \frac{q^*(1-p_s)}{1-q^* p_s} \delta \left(w_T(1) - w_T(0)\right)$$

■

Proof. (Proof of Corollary 1). (*i*) War occurs with positive probability if and only if (2) and (3) fail. Clearly, (3) is less stringent as p_s increases.

(*ii*) If (2) and (3) fail, the probability of preventive war is

$$1 - (1 - q^* p_s)\, r^* = 1 - \frac{\frac{k}{1-w_T(0)-w_D(0)}}{p_s + (1-p_s)\frac{\delta[w_T(1)-w_T(0)]}{1-w_T(0)-w_D(0)}}$$

which is decreasing in p_s.

(*iii*) If (2) and (3) fail, the share of preventive wars that are mistaken is

$$1 - \frac{(1 - r^*(1-p_s))\, q^*}{1-(1-q^* p_s)\, r^*} = 1 - \frac{1 - \frac{k}{\delta[w_T(1)-w_T(0)]}}{p_s + (1-p_s)\frac{\delta[w_T(1)-w_T(0)]}{1-w_T(0)-w_D(0)} - \frac{k}{1-w_T(0)-w_D(0)}}$$

which is decreasing in p_s. ■

2 Infinite-Horizon Game: Set-Up and Results

Consider the following extension of the baseline model. The game starts as in period 1 of the two-period game and each subsequent period repeats the same timing, as long as the target does not invest in military capabilities and peace prevails. If T invests at t, then it obtains additional military capabilities at $t+1$ only if D does not strike preventively. Once T obtains additional military capabilities, it conserves them and each period from then on proceeds as in period 2 of the two-period game. If T has its baseline level of military capabilities and suffers a preventive war at t, it loses the opportunity to militarize for N periods, *i.e.* the timing in periods $t+1$ to $t+N$ follows period 2 of the two-period game (with $M_{t+s} = 0 \ \forall 1 \leq s \leq N$), and the timing in period $t+N+1$ follows period

1 of the two-period game (and continues according to the same rules). We call N the *effectiveness* of a preventive war.[2] Finally, if T rejects an offer from D at t, then $t+1$ follows the same timing as t.

We analyze this game and impose that, at any information set, play is sequentially rational given beliefs, and beliefs are obtained using Bayes' rule whenever possible. There can be multiple equilibria in the infinite-horizon game. We first restrict attention to the set of *Markov Perfect Equilibria* (MPEs) of the game. An MPE requires that players play *Markovian* strategies, *i.e.,* strategies that they depend on history only through payoff-relevant state variables, here the military capabilities of the target (M_t), whether the target has the option to militarize (O_t) and if it does not, the number of periods (n_t) of the effectiveness of the preventive war that have passed already.[3] An MPE is a vector of Markovian strategies that are mutual best-responses, beginning at any date t for any value of the payoff-relevant state variables. For the full of solution of MPEs, see Proposition 7 in section 3.

We then consider the set of *Perfect Public Equilibria* (PPEs) of the game. A PPE requires that players play *public* strategies, *i.e.,* strategies that depend not just on the payoff-relevant state variables listed above, but on the full public history of the game, here the signals about T's militarization decision, the offers from D and the decisions by T to accept or reject D's offers. A PPE is a vector of public strategies that are mutual best-responses, beginning at any date t for any public history.

We ask whether countries can sustain the efficient outcome in a PPE, where militarization does not occur and peace prevails, if any public history revealing a deviation triggers the MPE. For simplicity, we restrict attention to the set of efficient PPEs in *stationary* strategies, where along the equilibrium path, T does not militarize, D makes a

[2]To be clear, a preventive war removes T's opportunity to militarize for N periods. Launching another preventive war before the N periods have passed does not extend the amount of time before T regains the opportunity to militarize. In other words, if T has the opportunity to militarize at t and D launches a preventive war at t, then T loses the opportunity to militarize from periods $t+1$ to $t+N$, and regains it in period $t+N+1$, whether or not D launched another preventive war between periods $t+1$ and $t+N$.

[3]$O_t \in (0,1)$, where $O_t = 1$ if and only if the target has the option to militarize and $n_t \in \{0, ..., N-1\}$, where we set $n_t = 0$ if $O_t = 1$.

fixed offer $z_t = z^*$ after $s_t = 0$, which T accepts. If there is no efficient PPE in stationary strategies, countries play the MPE.

In parallel with the two-period game, we characterize the threshold values for the effect of militarization. We say that the effect of militarization is smaller than the cost of investment, averaged over all periods, if

$$\delta\left(w_T\left(1\right) - w_T\left(0\right)\right) \leq \left(1-\delta\right)k \tag{4}$$

We say that the effect of militarization is smaller than the cost of preventive war, averaged over all periods, if

$$\delta\left(w_T\left(1\right) - w_T\left(0\right)\right) \leq \left(1-\delta\right)\left(1 - w_T\left(0\right) - w_D\left(0\right)\right) \tag{5}$$

We say that the effect of militarization is smaller than the cost of preventive war, averaged over the effectiveness of a preventive war, if

$$\delta\left(w_T\left(1\right) - w_T\left(0\right)\right) \leq \frac{1-\delta}{1-\delta^{N+1}}\left(1 - w_T\left(0\right) - w_D\left(0\right)\right) \tag{6}$$

We say that the signal is sufficienlty informative, or that the effect of militarization is smaller than the cost of investment, averaged over all periods, assuming it goes undetected, if

$$\left(1-p_s\right)\delta\left[w_T\left(1\right) - w_T\left(0\right)\right] \leq \left(1-\delta\right)k \tag{7}$$

First, we can show that an efficient MPE exists if militarization is not rationalizable or if it can be deterred. Militarization is not rationalizable if the effect of militarization is smaller than the cost of investment, averaged over all periods. Militarization can be deterred if the effect of militarization is greater than the cost of preventive war, averaged over all periods, and the signal is sufficiently informative. In either case, D need not fear T's militarization and peace prevails under the most favorable terms for D, $z^* = w_T\left(0\right)$.

Proposition 4 *(i) An efficient MPE exists if either (i.1) the effect of militarization is smaller than the cost of investment, averaged over all periods ((4) holds), or (i.2) the effect of militarization is greater than the cost of preventive war, averaged over all periods ((5) fails), and the signal is sufficiently informative ((7) holds).*

(ii) In such circumstances, an efficient PPE in stationary strategies always exists, where D offers $z^* = w_T(0)$ after signal $s_t = 0$.

Proof. *See section 3.* ■

Second, we note that there exists an MPE where peace prevails and militarization occurs if an investment in military capabilities is rationalizable and preventive war is too costly, i.e. if the effect of militarization is smaller than the cost of preventive war, averaged over the effectiveness of a preventive war. Now recall from the above proposition that if the effect of militarization is greater than the cost of preventive war, averaged over all periods, then it is possible to construct an efficient equilibrium if the signal is sufficiently informative. If D does not expect future militarization attempts, it would want to declare war in response to an unambiguous signal of militarization. Therefore, there is a unique MPE where peace prevails and militarization occurs if the investment is rationalizable and either preventive war is not rationalizable or it proves too costly, because of its limited effectiveness, and the signal is not sufficiently informative. In these cases, efficiency may not necessarily be ensured in a PPE in stationary strategies. As T becomes more patient, it demands greater - not smaller - concessions, which may be too high for D. [4] Formally:

Proposition 5 *(i) There is a unique MPE where peace prevails and T militarizes if the effect of militarization is greater than the cost of investment, averaged over all periods ((4) fails) and either (i.1) the effect of militarization is smaller than the cost of preventive war, averaged over all periods ((5) holds), or (i.2) the effect of militarization is greater than the cost of preventive war, averaged over all periods ((5) fails), but smaller than the cost of preventive war, averaged over the effectiveness of a preventive war ((6) holds), and the signal is not sufficiently informative ((7) fails).*

[4] Note that fixing $w_T(1) - w_T(0)$ and $1 - w_T(0) - w_D(0)$, then as δ approaches 1, condition (5) fails, but condition (6) may hold.

(ii) In such circumstances, an efficient PPE in stationary strategies, where D offers $z^* > w_T(0)$ after signal $s_t = 0$, exists if and only if

$$[1-\delta(p_s+(1-p_s)\delta)]\,\delta(w_T(1)-w_T(0)) \leq (1-\delta)k \tag{8}$$

This condition may fail even as countries become very patient (δ approaches one).

Proof. *See section 3.* ■

Finally, if the investment is rationalizable, preventive war is not too costly, and the signal is not sufficiently informative, then strategic uncertainty remains and war occurs with positive probability in the MPE.

We can hope that efficiency, and peace, can be sustained in a stationary PPE. In order to be dissuaded from militarization, T must receive concessions ($z^* > w_T(0)$). Yet concessions make preventive war attractive, as it ensures T's demilitarization without any concession. An efficient PPE in stationary strategies exists if the concessions needed to prevent militarization are smaller than the cost of a preventive war, averaged over the effectiveness of the preventive war (condition (11), below, holds).[5] This condition may fail even if countries are very patient, as explained in the text. Formally:

Proposition 6 *(i) There is a unique MPE[6] where war happens with positive probability if the effect of militarization is greater than the cost of investment, averaged over all periods ((4) fails), greater than the cost of preventive war, averaged over the effectiveness of a preventive war ((6) fails), and the signal is not sufficiently informative ((7) fails). In this MPE, T militarizes with probability*

$$q^* = \frac{1}{1+\frac{1-p_s}{1-\delta}\left[\frac{\delta(w_T(1)-w_T(0))}{\frac{1-\delta}{1-\delta^{N+1}}(1-w_T(0)-w_D(0))}-1\right]} \tag{9}$$

[5] The cost of war is discounted by δ since a deviation at t triggers war only at $t+1$.

[6] We ignore the possibility that D would respond to an unambiguous signal of a militarization attempt ($s_t = 1$) with a mixed strategy, which is not compelling. Knowing that T is investing in military capabilities, D should pick a pure strategy.

After a signal $s_t = 1$, preventive war happens with probability one. After a signal $s_t = 0$, peace prevails with probability

$$r^* = \frac{(1-\delta)\,k}{(1-p_s)\,\delta\,(w_T(1) - w_T(0))} \tag{10}$$

and preventive war happens with probability $1 - r^*$.

(ii) In such circumstances, an efficient PPE in stationary strategies, where D offers $z^* > w_T(0)$ after signal $s_t = 0$, exists if and only if

$$\frac{(1-p_s)\,\delta\,(w_T(1) - w_T(0)) - (1-\delta)\,k}{p_s + (1-p_s)\,\delta} \leq \delta \frac{1-\delta}{1-\delta^{N+1}}\,(1 - w_T(0) - w_D(0)) \tag{11}$$

This condition may fail even as countries become very patient (δ approaches one).

Proof. *See section 3.* ■

We now investigate the role of information problems and the effectiveness of preventive war on the likelihood of conflict. We conclude:

Corollary 3 *The greater is the informativeness of the signal*

(i) The more stringent become the conditions under which preventive wars happen with positive probability, and if such conditions are met,

(ii) The smaller is the probability of preventive war;

(iii) The smaller is the share of preventive wars that are mistaken.

Proof. *See section 3.* ■

Next, we investigate the effect of the effectiveness of preventive war on the likelihood of war. The statement is summarized in Corollary 2 in the main text (for a proof, see section 3).

3 Infinite-Horizon Game: Proofs

We first provide a characterization of the MPEs of this game, and then provide a proof of the formal statements presented in the previous section.

Proposition 7 *(A) In a period t where T does not have the option to militarize, there is a unique MPE: D offers $z_t^* = w_T(M_t)$ and T accepts $z_t \geq w_T(M_t)$.*

(B) Consider a period t where T has the option to militarize.

(B.i) In an efficient MPE: T chooses $q^* = 0$. After $s_t = 0$, D offers $z^* = w_T(0)$. After $s_t = 1$, D offers $z^* = w_T(0)$ if (5) holds and declares war otherwise. T accepts $z_t \geq w_T(0)$.

An efficient MPE exists if and only if either

(B.i.1) (5) and (4) hold,

(B.i.2) (5) fails and (7) holds.

(B.ii) In an inefficient MPE where peace prevails: T chooses $q^* = 1$. After any s_t, D offers $z^* = w_T(0)$. T accepts $z_t \geq w_T(0)$. This MPE exists if (4) fails, (6) holds.

(B.iii) In an inefficient MPE where war happens with positive probability, T chooses q^* given by (9). After $s_t = 1$, D declares war. After $s_t = 0$, D offers $z_t^* = w_T(0)$ with probability r^* given by (10) and declares war with probability $1-r^*$. T accepts $z_t \geq w_T(0)$. This MPE exists if neither (6) nor (7) hold.

Proof. *First, T accepts any $z_t \geq w_T(M_t)$, since strategies are not history-dependent.*

Now consider (A). If $M_t = 1$, D offers $z^* = w_T(1)$, since war is inefficient and $1 - w_T(1)$ is from then on D's maximum per-period payoff. Likewise, if $M_t = 0$, D offers $z^* = w_T(0)$, since war is inefficient and $1 - w_T(0)$ is D's maximum per-period payoff.

Now consider (B). After any signal, D chooses between declaring war and offering $z_t = w_T(0)$.

After $s_t = 1$, D declares war if and only if

$$w_D(0) + \delta\left(1-\delta^N\right)\frac{1-w_T(0)}{1-\delta} + \delta^{N+1}V_D^{MPE}(0,1,0) > 1 - w_T(0) + \delta\frac{1-w_T(1)}{1-\delta} \tag{12}$$

$$\Leftrightarrow \delta\left(w_T(1) - w_T(0)\right) - (1-\delta)\left(1 - w_T(0) - w_D(0)\right) > \delta^{N+1}\left[1 - w_T(0) - (1-\delta)V_D^{MPE}(0,1,0)\right] \tag{13}$$

where $V_i^{MPE}(M_t, O_t, n_t)$ is country i's continuation value in the MPE.

After $s_t = 0$, D declares war if and only if

$$\begin{aligned} &w_D(0) + \delta\left(1-\delta^N\right)\frac{1-w_T(0)}{1-\delta} + \delta^{N+1}V_D^{MPE}(0,1,0) > \\ &\qquad 1 - w_T(0) + \delta\frac{q^*(1-p_s)}{1-q^*p_s}\frac{1-w_T(1)}{1-\delta} + \delta\frac{1-q^*}{1-q^*p_s}V_D^{MPE}(0,1,0) \end{aligned} \tag{14}$$

$$\begin{aligned} &\Leftrightarrow \delta\left(w_T(1) - w_T(0)\right) - (1-\delta)\left(1 - w_T(0) - w_D(0)\right) > \\ &\delta^{N+1}\left[1 - w_T(0) - (1-\delta)V_D^{MPE}(0,1,0)\right] + \delta\frac{1-q^*}{1-q^*p_s}\left((1-\delta)V_D^{MPE}(0,1,0) - (1-w_T(1))\right) \end{aligned} \tag{15}$$

Clearly, if (15) holds, then so does (13) (since $V_D^{MPE}(0,1,0) \geq \frac{1-w_T(1)}{1-\delta}$). Thus either (i) D offers $z^* = w_T(0)$ after any s_t, or (ii) D declares war after $s_t = 1$ and, after $s_t = 0$, D offers $z_t = w_T(0)$ with probability $r^* \in [0,1]$ and declares war with probability $1 - r^*$.[7] *In (i), T prefers $I_t = 1$ if $-k + w_T(0) + \delta\frac{w_T(1)}{1-\delta} > w_T(0) + \delta V_T^{MPE}(0,1,0)$, or*

$$k < \delta\left[\frac{w_T(1)}{1-\delta} - V_T^{MPE}(0,1,0)\right] \tag{16}$$

[7] Here, we ignore the possibility that D would respond to an unambiguous signal of a militarization attempt ($s_t = 1$) with a mixed strategy, which is not compelling. Knowing that T is investing in military capabilities, D should pick a pure strategy.

In (ii), T prefers $I_t = 1$ if

$$-k + w_T(0) + \delta(1-(1-p_s)r^*)V_T^{MPE}(0,0,0) + \delta(1-p_s)r^*\frac{w_T(1)}{1-\delta} > w_T(0) + \delta(1-r^*)V_T^{MPE}(0,0,0) + \delta r^* V_T^{MPE}(0,1,0)$$

$$\Leftrightarrow k < \delta r^*\left[(1-p_s)\left(\frac{w_T(1)}{1-\delta} - V_T^{MPE}(0,1,0)\right) - p_s\left(1-\delta^N\right)\left(V_T^{MPE}(0,1,0) - \frac{w_T(0)}{1-\delta}\right)\right] \tag{17}$$

using the fact that the value of the game for T after a preventive strike at t is

$$V_T^{MPE}(0,0,0) = \frac{1-\delta^N}{1-\delta}w_T(0) + \delta^N V_T^{MPE}(0,1,0) \tag{18}$$

Let us characterize the conditions under which an efficient MPE exists. In an efficient MPE, $V_D^{MPE}(0,1,0) = \frac{1-w_T(0)}{1-\delta}$, $V_T^{MPE}(0,1,0) = \frac{w_T(0)}{1-\delta}$. From (15), D offers $z^* = w_T(0)$ after $s_t = 0$ since war is inefficient. From (13), D offers $z^* = w_T(0)$ after $s_t = 1$ if (5) holds and declares war otherwise. From (16) and (17), T prefers not to invest if and only if either (i) (5) and (4) hold or (ii) (5) fails and (7) holds.

Now assume that the efficient equilibrium does not exist and peace prevails. Thus $q^* > 0$ and D offers $z^* = w_T(0)$ after any s_t. Generically, we cannot have $q^* \in (0,1)$.[8] *Therefore,*

$$V_D^{MPE}(0,1,0) = 1 - w_T(0) + \frac{\delta}{1-\delta}(1 - w_T(1)) \tag{19}$$

$$V_T^{MPE}(0,1,0) = w_T(0) - k + \frac{\delta w_T(1)}{1-\delta} \tag{20}$$

Using (13), D offers $z^* = w_T(0)$ after any s_t if and only if (6) holds. Using (16), T chooses $q^* = 1$ if and only if holds (4) fails.

[8] In non-generic regions of the parameter space, any player breaks indifference in favor of the efficient action. Thus, if D plays a pure strategy after any signal and T is indifferent about investing, then $q^* = 0$.

Now assume that the efficient equilibrium does not exist and war occurs with positive probability. D must declare war after $s_t = 1$ and, after $s_t = 0$, offer $z_t = w_T(0)$ with probability $r^* \in [0,1]$ and declare war with probability $1 - r^*$. Now note that $q^* > 0$. Indeed, if $q^* = 0$ then $r^* = 1$ ((14) fails, since war is inefficient), so that war does not occur, a contradiction. Second, note that $r^* > 0$. Indeed, $r^* = 0$ implies $q^* = 0$ ((17) fails), which we just ruled out. Third, note that $q^* < 1$. Indeed if $q^* = 1$, then (13) and (15) are equivalent, so that $r^* = 0$, which we just ruled out. Next, $q^* \in (0,1)$ implies that T is indifferent about investing, i.e. $V_T^{MPE}(0,1,0) = w_T(0) + \delta(1-r^*) V_T^{MPE}(0,0,0) + \delta r^* V_T^{MPE}(0,1,0)$, or

$$V_T^{MPE}(0,1,0) = \frac{w_T(0) - \delta(1-r^*)\left(V_T^{MPE}(0,1,0) - V_T^{MPE}(0,0,0)\right)}{1-\delta} \tag{21}$$

Rearranging (18), we get

$$V_T^{MPE}(0,1,0) - V_T^{MPE}(0,0,0) = \left(1-\delta^N\right)\left[V_T^{MPE}(0,1,0) - \frac{w_T(0)}{1-\delta}\right] \tag{22}$$

(21) and (22) imply $V_T^{MPE}(0,1,0) = V_T^{MPE}(0,0,0) = \frac{w_T(0)}{1-\delta}$. Replacing in (17), which must hold with equality, we solve for r^* and obtain (10). Generically, we cannot have $r^* = 1$.[9] *Moreover, $r^* < 1$ if and only if (7) fails. Now since D declares war after $s_t = 1$, and is indifferent about declaring war after $s_t = 0$, we get $V_D^{MPE}(0,1,0) = w_D(0) + \delta\left(1-\delta^N\right)\frac{1-w_T(0)}{1-\delta} + \delta^{N+1} V_D^{MPE}(0,1,0)$ or*

$$V_D^{MPE}(0,1,0) = \frac{1}{1-\delta}\left[1 - w_T(0) - \frac{1-\delta}{1-\delta^{N+1}}\left(1 - w_T(0) - w_D(0)\right)\right] \tag{23}$$

Replacing in (15), which must hold with equality, we solve for q^* and obtain (9). $q^* \in (0,1)$ if and only if (6) fails. ■

Proof. (Proof of Proposition 4) Part (i) follows from proposition 7. Part (ii) is straightforward since Markovian strategies are stationary and depend trivially on history. ■

[9] In non-generic regions of the parameter space, any player breaks indifference in favor of the efficient action. If D plays a pure strategy after any signal and T is indifferent about investing, then $q^* = 0$.

Proof. (Proof of Proposition 5) Part (i) follows from proposition 7. For part (ii), recall that $V_D^{MPE}(0,1,0)$ and $V_T^{MPE}(0,1,0)$ are given by (19) and (20) respectively. Let us construct an efficient PPE in stationary strategies.

First, T must accept z^*, *i.e.* $w_T(0) + \delta V_T^{MPE}(0,1,0) \leq \frac{z^*}{1-\delta}$, or

$$z^* \geq w_T(0) + \delta\left[\delta\left(w_T(1) - w_T(0)\right) - (1-\delta)k\right] \tag{24}$$

Next, after any history revealing a deviation, T accepts any $z_t \geq w_T(0)$, since from $t+1$ countries play the MPE.

Moving up, D offers $z_t = w_T(0)$ after $s_t = 1$ ((13) fails since (6) holds).

Next, T refrains from militarization if and only if

$$-k + p_s w_T(0) + (1-p_s) z^* + \delta \frac{w_T(1)}{1-\delta} \leq \frac{z^*}{1-\delta} \tag{25}$$

$$\Leftrightarrow z^* \geq w_T(0) + \frac{\delta\left(w_T(1) - w_T(0)\right) - (1-\delta)k}{p_s + (1-p_s)\delta} \tag{26}$$

which is a tighter condition than (24).

Also, the best deviation for D, after $s_t = 0$, is to offer $z_t = w_T(0)$ ((15) fails since (6) holds). D does not offer $z_t = w_T(0)$ if and only if $1 - w_T(0) + \delta V_D^{MPE}(0,1,0) \leq \frac{1-z^*}{1-\delta}$, or

$$z^* \leq w_T(0) + \delta^2\left(w_T(1) - w_T(0)\right) \tag{27}$$

(26) and (27) hold if and only if (8) holds. (6) implies (8) if and only if

$$\frac{1-\delta}{1-\delta^{N+1}}\left(1 - w_T(0) - w_D(0)\right) \leq \frac{(1-\delta)k}{1-\delta\left(p_s + (1-p_s)\delta\right)} \tag{28}$$

Taking the limit as δ approaches 1, and using l'Hopital's rule, this becomes

$$\frac{1}{N+1}\left(1 - w_T(0) - w_D(0)\right) \leq \frac{k}{2-p_s} \tag{29}$$

which may fail. If $1 - w_T(0) - w_D(0) = (N+1)\frac{k}{1-p_s} = (N+2)(w_T(1) - w_T(0))$, then $\exists\delta' \in (0,1)$ such that $\forall\delta \in (\delta', 1)$, (4) fails, (6) holds and yet (8) fails. ■

Proof. (Proof of Proposition 6). Part (i) follows from proposition 7. For part (ii), recall that $V_D^{MPE}(0,1,0)$ is given by (23) and $V_T^{MPE}(0,1,0) = \frac{w_T(0)}{1-\delta}$. Let us construct an efficient PPE in stationary strategies.

First, T must accept z^*, *i.e.* $w_T(0) + \delta V_T^{MPE}(0,1,0) \leq \frac{z^*}{1-\delta}$, or $z^* \geq w_T(0)$. Next, after any history revealing a deviation, T accepts any $z_t \geq w_T(0)$, since from $t+1$ countries play the MPE.

Moving up, D must declare war after $s_t = 1$ ((13) holds since (6) fails).

Next, T does not want militarize if and only if

$$-k + p_s \frac{w_T(0)}{1-\delta} + (1-p_s)\left(z^* + \delta \frac{w_T(1)}{1-\delta}\right) \leq \frac{z^*}{1-\delta} \tag{30}$$

$$\Leftrightarrow z^* \geq w_T(0) + \frac{(1-p_s)\,\delta\,(w_T(1) - w_T(0)) - (1-\delta)\,k}{p_s + (1-p_s)\,\delta} \tag{31}$$

which ensures that T accepts $z^* \geq w_T(0)$.

Third, the best deviation for D, after $s_t = 0$, is to offer $z_t = w_T(0)$, since it gives $1 - w_T(0) + \delta V_D^{MPE}(0,1,0)$, strictly greater than $V_D^{MPE}(0,1,0)$, the payoff of declaring war. D does not offer $z_t = w_T(0)$ if and only if

$$1 - w_T(0) + \delta V_D^{MPE}(0,1,0) \leq \frac{1-z^*}{1-\delta} \tag{32}$$

$$\Leftrightarrow z^* \leq w_T(0) + \delta \frac{1-\delta}{1-\delta^{N+1}}\,(1 - w_T(0) - w_D(0)) \tag{33}$$

(31) and (33) hold if and only if (11) holds. (11) may fail. For example, if $w_T(1) - w_T(0) = f\,(1 - w_T(0) - w_D(0))$, for $f > \frac{1}{(N+1)(1-p_s)}$, then $\exists\delta'' \in (0,1)$ such that $\forall\delta \in (\delta'', 1)$, (6) and (7) fail and yet (11) fails.[10] ■

Proof. (Proof of Corollary 3). War occurs with positive probability if and only if (6), (7), and (11) fail.

[10]We need not impose any restriction on k to obtain this result, and in fact we want to maintain the plausible assumption that the cost of investment is less than the cost of preventive war ($k < 1 - w_T(0) - w_D(0)$).

(i) As p_s increases, (7) and (11) are less stringent.

(ii) If (6), (7), and (11) fail, the probability of preventive war is

$$1-(1-q^*p_s)\,r^* = 1-\frac{(1-\delta)\,k}{\delta\,(w_T\,(1)-w_T\,(0))}\frac{1+\frac{1}{1-\delta}\left(\frac{\delta(w_T(1)-w_T(0))}{\frac{1-\delta}{1-\delta^{N+1}}(1-w_T(0)-w_D(0))}-1\right)}{1+\frac{1-p_s}{1-\delta}\left(\frac{\delta(w_T(1)-w_T(0))}{\frac{1-\delta}{1-\delta^{N+1}}(1-w_T(0)-w_D(0))}-1\right)} \tag{34}$$

which is decreasing in p_s.

(iii) If (6), (7), and (11) fail, the share of preventive wars that are mistaken is

$$1-\frac{(1-r^*\,(1-p_s))\,q^*}{1-(1-q^*p_s)\,r^*} = 1-\frac{1-\frac{(1-\delta)k}{\delta[w_T(1)-w_T(0)]}}{\frac{1-(1-q^*p_s)r^*}{q^*}}$$

which is decreasing in p_s if $\partial\frac{\frac{1-(1-q^*p_s)r^*}{q^*}}{\partial p_s}<0$. This is indeed the case since, using (34),

$$\begin{aligned}\frac{1-(1-q^*p_s)\,r^*}{q^*} &= 1+\frac{1-p_s}{1-\delta}\left(\frac{\delta\,(w_T\,(1)-w_T\,(0))}{\frac{1-\delta}{1-\delta^{N+1}}\,(1-w_T\,(0)-w_D\,(0))}-1\right)\\ &\quad -\frac{(1-\delta)\,k}{\delta\,(w_T\,(1)-w_T\,(0))}\left(1+\frac{1}{1-\delta}\left(\frac{\delta\,(w_T\,(1)-w_T\,(0))}{\frac{1-\delta}{1-\delta^{N+1}}\,(1-w_T\,(0)-w_D\,(0))}-1\right)\right)\end{aligned}$$

■

Proof. (Proof of Corollary 2). War occurs with positive probability if and only if (6), (7), and (11) fail.

(i) As N increases, (6) and (11) are more stringent.

(ii) The probability of preventive war, $1-(1-q^*p_s)\,r^*$, decreases in N since q^* decreases in N.

(iii) The share of preventive wars that are mistaken increases in N since $\frac{1-r^*}{q^*}+r^*p_s$ increases in N, given that q^* decreases in N. ■