Defensive Weapons and Defensive Alliances

Sylvain Chassang — Princeton University

Gerard Padró i Miquel* — London School of Economics and NBER

December 17, 2008

In 2002, U.S. President George W. Bush initiated the deployment of a new ballistic missile defense system.[1] The move triggered vociferous international concerns, including a recent statement of Russia and China condemning U.S. plans as a destabilizing move.[2] Indeed, the move amounts to a withdrawal from the 1972 Anti-Ballistic Missile treaty. The U.S. position is that such missile defense systems only reduce the damage caused by an incoming strike, and therefore do not threaten international stability. This paper provides a careful formal analysis of how the unilateral acquisition of defensive weapons may affect the sustainability of peace.

We consider a dynamic game in which two symmetric countries repeatedly decide to be peaceful or to attack. Peace is sustained in equilibrium by trigger strategies in which attacks are followed by permanent conflict. Under complete information, peace is sustainable if and only if the value of continued peace is greater than the temptation of launching a surprise attack. Because defensive weapons limit the possibility of retaliation, the unilateral acquisi-

*Chassang: Department of Economics, Princeton University, Princeton, NJ 08544, USA (email: chassang@princeton.edu). Padró i Miquel: London School of Economics, Houghton Street, London WC2A 2AE, UK (email: g.padro@lse.ac.uk).

We thank Micael Castanheira for suggesting the analysis of defensive alliances. Any errors are, of course, our own.

[1] National Security Presidential Directive 23, December 16, 2002.

[2] The statement was issued in Beijing in May 23, 2008, during a diplomatic visit by Russian President Dmitry Medvedev.

tion of defensive weapons raises the stronger country's predatory incentives and reduces the sustainability of peace.

We then turn to the more nuanced model of Chassang and Padró i Miquel (2008) in which fear, rather than just greed, can be a motive for conflict. In that setting, peace is sustainable only if both predatory and preemptive incentives are small enough. Our theoretical results in this setting uphold the intuition that by reducing the damage caused by surprise attacks, defensive weapons can decrease the need to launch preemptive strikes, and improve the sustainability of peace. However, since defensive weapons also increase predatory incentives in this setting, the net effect is ambiguous. Careful analysis suggests that the net effect will be destabilizing when players are patient. Indeed, defensive weapons reduce preemptive incentives only in the current period, whereas they increase predatory incentives in all future periods. Hence, there are important reasons to suspect that the unilateral acquisition of defensive weapons undermines the prospects of peace even under strategic risk.

The main reason for this result is that defensive weapons protect against aggressive surprise attacks but also against rightful retaliation. In contrast, we show that defensive alliances which distinguish between different types of conflict can unambiguously increase the sustainability of peace. In particular, treaties by which countries agree to help each other only when they suffer an unprovoked attack reduce preemptive incentives without increasing predatory incentives. Hence, a country's unilateral decision to join such a defensive alliance will not be destabilizing. NATO is a prominent example of such defensive alliances.

The paper is organized as follows: Section 1 describes the framework; Section 2 studies the effect of defensive weapons under complete information; Section 3 introduces strategic risk and explores how it changes predictions; Section 4 turns to defensive alliances; Section 5 concludes. Proofs are contained in the appendix.

1 A Simple Model of Peace and Conflict

We consider two countries $i \in \{1,2\}$ that play an infinite horizon repeated game, with discrete time $t \in \mathbb{N}$, and share a common discount factor δ. Each period t, the players simultaneously decide whether to be peaceful (P) or to attack (A). The stage game payoffs are as follows:

	P	A
P	π	$s(d_i, d_{-i})$
A	$f(d_i, d_{-i})$	$w(d_i, d_{-i})$

where payoffs are given for row player i. Parameter π represents flow payoffs from peace, w corresponds to the payoff of simultaneous conflict, f and s respectively correspond to the payoffs of being first and second mover during a surprise attack. Finally, $d_i \in \{0,1\}$ indicates whether player i has acquired defensive weapons or not. Note that in this paper, d_i is a parameter of the game that is taken as given. It is not a decision variable. Let us denote this complete information game by Γ_{CI}. We make the following assumption.

Assumption 1 (First Strike Advantage) $\forall (d_i, d_{-i}) \in \{0,1\}^2$, $f(d_i, d_{-i}) > w(d_i, d_{-i}) > s(d_i, d_{-i})$.

Assumption 1 implies that there is a first mover advantage. As the payoffs from a simultaneous attack dominate those from being a second mover, there is a preemptive motive for war. The acquisition of defensive weapon, such as missile defense systems or bunkers, affects payoffs as follows:

$$f(d_i, d_{-i}) = f_0 - d_{-i}\mu_f \; ; \;\; w(d_i, d_{-i}) = w_0 + d_i\mu_w \; ; \;\; \text{and} \;\; s(d_i, d_{-i}) = s_0 + d_i\mu_s$$

where μ_f, μ_w and μ_s are positive constants. In words, acquiring defensive weapons decreases the payoffs of a possible first mover aggressor, while increasing one's payoffs in the event of simultaneous or second mover conflict.

We focus exclusively on *trigger strategies*, such that whenever a player attacks, then both players attack in every subsequent period. Assumption 1 implies attacking in every period is indeed a subgame perfect equilibrium. We denote by $W(d_i, d_{-i}) \equiv \sum_{t=0}^{+\infty} \delta^t w(d_i, d_{-i})$, $F(d_i, d_{-i}) \equiv f(d_i, d_{-i}) + \delta W(d_i, d_{-i})$ and $S(d_i, d_{-i}) \equiv s(d_i, d_{-i}) + \delta W(d_i, d_{-i})$, the equilibrium values that players obtain upon simultaneous, first mover and second mover attacks.

Throughout the paper, we consider the case where player $-i$ does not have defensive weapons ($d_{-i} = 0$) and explore how the acquisition of weapons by player i ($d_i \in \{0, 1\}$) affects the sustainability of peace.

2 Defensive Weapons under Complete Information

In the complete information game Γ_{CI} peace is sustainable if and only if predatory incentives are low enough, i.e. if the value of continued cooperation is greater than the deviation temptation. This corresponds to the following proposition.

Proposition 1 (Peace under Complete Information) *Perpetual peace is sustainable in equilibrium in game Γ_{CI} if an only if*

$$\forall i \in \{1, 2\}, \quad \frac{1}{1-\delta}\pi > F(d_i, d_{-i}). \tag{1}$$

Note that the difference $F(d_i, d_{-i}) - \pi/(1-\delta)$ corresponds to what a player gains from attacking a peaceful opponent, which we call the predatory incentives of player i. Equation (1) highlights that under complete information, peace is sustainable if and only if predatory incentives of both players are low enough. Note that payoffs S and W do not matter in determining whether or not peace is sustainable. In contrast, these payoffs will play an important role in Section 3 where we introduce a preemptive motive for conflict.

Let us denote by π_{CI} the smallest value of π such that peace is sustainable under complete information. It follows from condition (1) that $\pi_{CI} = (1-\delta)\max_i F(d_i, d_{-i})$. The higher π_{CI}, the more difficult it is to sustain peace under complete information. The question is whether defensive weapons facilitate sustaining peace.

Proposition 2 (Defensive Weapons under Complete Information) *Under complete information, the unilateral acquisition of defensive weapons reduces the sustainability of peace. Formally, we have that* $\pi_{CI}(1,0) > \pi_{CI}(0,0)$*.*

This result follows from the fact that defensive weapons increase predatory incentives, $F - \pi/1-\delta$. To gain intuition, it is useful to divide the total payoffs F obtained from a predatory attack into two parts. First, there is the immediate gain from the surprise strike, $f(d_i, d_{-i})$. Second, this is followed by retaliatory conflict, which gives a value $\frac{\delta}{1-\delta}w(d_i, d_{-i})$ to the attacker. The unilateral acquisition of weapons does not change immediate gains from attack, but it increases the payoffs obtained in the subsequent conflict by an amount $\frac{\delta}{1-\delta}\mu_w$. This increases the temptation to attack. In short, this problem arises because defensive weapons that can shield a country against an aggressive surprise attack, can also shield an aggressor against righteous retaliation from the victim. Since the threat of retaliation is necessary to deter predatory behavior, the unilateral acquisition of defensive weapons is destabilizing. This is related to Powell (2003), which shows that in the context of crisis bargaining, defensive weapons can have negative effects by increasing the assertiveness of whoever owns them.

3 Defensive Weapons Under Strategic Risk

The previous section shows that under complete information, defensive weapons increase predatory incentives and are destabilizing. Intuitively, this ignores an important benefit of defensive weapons, which is to reassure whoever owns them. When fear is a motive for conflict, war might be initiated by a preemptive strike launched because the attackers fear suffering a predatory strike. Since a country's defensive weaponry shields it from predatory strikes, it may reduce its incentives to launch preemptive strikes and thereby increase stability. Under the complete information setting, this effect does not occur because players perfectly anticipate each other's moves and coordination presents no difficulty. This section shows that once strategic risk is taken into account as in Chassang and Padró i Miquel (2008), the unilateral acquisition of weapons may indeed ease the sustainability of peace.

We use the results of Chassang and Padró i Miquel (2008) without proof, but provide a description of their modelling approach. Strategic risk is modelled by introducing the idea that players make private noisy assessments of their environment. This creates a possibility of miscoordination in equilibrium which pushes players to second guess each other's decisions. This introduces fear as a motive for conflict and considerably restricts the scope for cooperation, even when players get very precise information about their environment. Formally, we consider a setting where the payoffs π from peace are in fact an i.i.d. random sequence $(\tilde{\pi}_t)_{t\in\mathbb{N}}$, distributed according to an integrable distribution g with support $(-\infty, +\infty)$. In each period t, the players' stage game payoffs are

	P	A
P	$\tilde{\pi}_t$	$s(d_i, d_{-i})$
A	$f(d_i, d_{-i})$	$w(d_i, d_{-i})$.

In contrast to the complete information setting, the flow payoffs of peace $\tilde{\pi}_t$ are not directly observable by the players at the time of decision. Instead, each player $i \in \{1,2\}$ observes a private signal of the form $x_{i,t} = \tilde{\pi}_t + \sigma\epsilon_{i,t}$ where $\{\epsilon_{i,t}\}_{i\in\{1,2\},\, t\in\mathbb{N}}$ is an i.i.d. sequence of centered errors with support $[-1,1]$. This corresponds to a global games information structure à la Carlsson and van Damme (1993). For simplicity we assume that $\tilde{\pi}_t$ is observable in period $t+1$ via the flow payoffs. Let us denote this game by $\Gamma_{\sigma,g}$. In order to draw meaningful comparisons with the complete information setting, we are interested in the properties of game $\Gamma_{\sigma,g}$ when its payoffs and information structure become arbitrarily close to those of the complete information game Γ_{CI}. For this purpose we study the properties of $\Gamma_{\sigma,g}$ as first σ goes to 0 and then g converges to d_π, the unit mass at π.

Chassang and Padró i Miquel (2008) show that as σ goes to 0 and g converges to the unit mass d_π, peace is sustainable in an equilibrium of game $\Gamma_{\sigma,g}$ if and only if the following condition holds:[3]

$$\prod_{i\in\{1,2\}} \left(\frac{1}{1-\delta}\pi - F(d_i, d_{-i})\right)^+ > \prod_{i\in\{1,2\}} \left(W(d_i, d_{-i}) - S(d_i, d_{-i})\right) \tag{2}$$

where $x^+ = \max(x, 0)$ for all $x \in \mathbb{R}$. Condition (2) corresponds to (P,P) being risk-dominant in the one-shot two-by-two game,

	P	A
P	Π	$S(d_i, d_{-i})$
A	$F(d_i, d_{-i})$	$W(d_i, d_{-i})$

where $\Pi = \frac{1}{1-\delta}\pi$. In contrast, the complete information condition (1) corresponds to (P,P) being a Nash equilibrium of the same game. Obviously, condition (2) is strictly more restrictive. In particular, peace is sustainable only if both predatory incentives $F - \Pi$, and

[3] Note that Chassang and Padró i Miquel (2008) consider the slightly different framework of exit games. It is immediate to show that trigger strategies of a repeated game map into equilibria of an appropriately chosen exit game.

preemptive incentives $W - S$ are low enough.

Let us define $\pi_{SU}(d_i, d_{-i})$ as the smallest value of π such that condition (2) holds. We clearly have that $\pi_{SU}(d_i, d_{-i}) > \pi_{CI}(d_i, d_{-i})$. The question is how the unilateral acquisition of weapons affects the sustainability of peace under strategic risk. Let us define

$$\overline{\mu}_s \equiv \frac{1}{1-\delta}\mu_w - \mu_f + \frac{\delta}{1-\delta}\frac{\mu_w \mu_f}{w-s}. \tag{3}$$

Proposition 3 (Defensive Weapons under Strategic Risk) *Whenever $\mu_s > \overline{\mu}_s$, then the unilateral acquisition of defensive weapons facilitates the sustainability of peace, i.e. $\pi_{SU}(1,0) < \pi_{SU}(0,0)$. Whenever $\mu_s < \overline{\mu}_s$, then the unilateral acquisition of defensive weapons reduces the sustainability of peace, i.e. $\pi_{SU}(1,0) > \pi_{SU}(0,0)$.*

Proposition 3 highlights that when fear is a motive for conflict, the unilateral acquisition of weapons can facilitate the sustainability of peace. This occurs when μ_s is high so that defensive weapons are particularly helpful to the victim of a surprise attack. In such circumstances, defensive weapons reduce preemptive incentives for conflict and this effect more than compensates the increase in predatory incentives. Hence, under strategic risk, it is possible for the unilateral acquisition of defensive weapons to improve the sustainability of peace.

However, as δ approaches 1, $\overline{\mu}_s$ becomes arbitrarily large. Therefore, for sufficiently patient players, defensive weapons will be destabilizing even under strategic risk. To understand why this is the case, we unbundle the different effects of defensive weapons under strategic risk. First, as under complete information, acquiring defensive weapons increases one's predatory incentives $F - \pi/(1-\delta)$ by an amount $\frac{\delta}{1-\delta}\mu_w$, corresponding to increased payoffs during retaliatory conflict. Second, acquiring defensive weapons reduces one's preemptive incentives by an amount $\mu_s - \mu_w$:

$$W(1,0) - S(1,0) = W(0,0) - S(0,0) - (\mu_s - \mu_w).$$

Note that this reduction in preemptive incentives corresponds entirely to changes in payoffs during the first period of conflict. In later periods, payoffs are increased by an amount μ_w independently of how conflict started and this change has no further effect on preemptive incentives.[4] As δ approaches 1, the increase in future predatory payoffs dominates the reduction in current preemptive incentives. Hence, when players are patient, the unilateral acquisition of defensive weapons is destabilizing, unless these defensive weapons are disproportionately more helpful to the victim of a surprise attack than during simultaneous conflict. While this may well be the case in some settings, we believe that this significantly weakens the case for defensive weapons to be stabilizing, even when fear is a motive for conflict.

4 Defensive Alliances

Under strategic risk, it is possible for defensive weapons to be stabilizing, but there is no guarantee that this is true. The main reason why defensive weapons need not be stabilizing is that they affect all payoffs F, W and S simultaneously. Since this protective effect is not conditional on who started the war, defensive weapons increase predatory incentives and do not diminish preemptive incentives as much as casual intuition would suggest.

To be unambiguously stabilizing, an ideal unilateral security arrangement should increase second mover payoffs S and decrease first and simultaneous mover payoffs F and W. This decreases both preemptive and predatory incentives. Defensive weapons do not achieve this because they are unconditionally helpful. In contrast, defensive alliances with third

[4] A third effect corresponds to the fact that acquiring weapons reduces the predatory incentives of one's opponent by an amount μ_f. Under strategic risk this also reduces one's opponent propensity to launch preemptive attacks and facilitates the sustainability of peace.

parties may afford greater degrees of freedom. In particular, to the extent that actions are observable to third parties, defensive alliances can be conditioned on the particular circumstances leading to conflict, and tailored to increase only the value S of suffering a surprise attack. In this section, we study the effect of a unilateral alliance such that one country receives foreign military assistance, but only in the case where it is attacked first. We show that unilaterally entering such an agreement (with an unmodeled third party) unambiguously improves the sustainability of peace.

Consider a situation where $d_i = d_{-i} = 0$. Assume that when country i signs a defensive alliance with a third country, this third country fights on i's side if and only if i suffers a surprise attack. Consequently, after the initial attack, flow payoffs to the victim are $w_0 + \eta$ instead of w_0, while flow payoffs to the aggressor are $w_0 - \eta$ instead of w_0, with $\eta > 0$. Denote by $s_i \in \{0, 1\}$ whether or not player i belongs to defensive alliance. Thresholds $\pi_{SU}(s_i, s_{-i})$ and $\pi_{CI}(s_i, s_{-i})$ are defined as in the previous section. The following proposition holds.

Proposition 4 (Defensive Alliances) *Under complete information, unilateral membership in a defensive alliance does not affect the sustainability of peace:* $\pi_{CI}(1, 0) = \pi_{CI}(0, 0)$.

Under strategic uncertainty, unilateral membership in a defensive alliance strictly improves the sustainability of peace: $\pi_{SU}(1, 0) < \pi_{SU}(0, 0)$.

Because the defensive alliance can be conditioned on the circumstances leading to conflict, it decreases preemptive incentives without increasing predatory incentives. An obvious concern is the observability of who attacked first. If actions are unobservable, entering a defensive alliances is not different from acquiring defensive weapons, as it improves payoffs unconditionally upon conflict. In such circumstances, defensive alliances may be destabilizing.

5 Conclusion

This paper explores how the unilateral acquisition of defensive weapons affects the sustainability of peace in a context with two symmetric countries. As in Chassang and Padró i Miquel (2008), we find that carefully taking into account strategic risk substantially changes the analysis. In particular, while defensive weapons are destabilizing under complete information, they can be stabilizing under strategic risk. However, this holds only when the players are impatient, or when defensive weapons are disproportionately more useful during a surprise attack than during a simultaneous or a first mover conflict.

Our analysis suggests that conditional defensive alliances are more likely to unambiguously improve the sustainability of peace. However, this holds only if actions are observable and one can determine who initiated conflict. This is a potential issue with the extension of NATO, and has been an important concern in the relationship between the U.S. and Taiwan. In particular, the 1979 Taiwan Relations Act, which replaced a 1955 mutual defense treaty, explicitly avoided a collective defense clause.

The simple analysis we pursue here leaves open a number of questions. First, one might want to extend the analysis to asymmetric countries. Another question is, what occurs when weapon stocks are endogenous?

A Proofs

Proof of Proposition 3: Tedious but simple algebra shows that

$$\begin{aligned}\frac{2}{1-\delta}\left[\pi_{SU}(1,0)-\pi_{SU}(0,0)\right] &= \frac{\delta}{1-\delta}\mu_W-\mu_f-2(w_0-s_0)\\ &+ \sqrt{\left(\frac{\delta}{1-\delta}\mu_W+\mu_f\right)^2+4(w_0-s_0)(w_0-s_0+\mu-\mu_s)}.\end{aligned}$$

This expression is decreasing in μ_s. Value $\overline{\mu}_s$ solves equation $\pi_{SU}(1,0)-\pi_{SU}(0,0)=0$. ■

References

[1] Carlsson, Hans and Eric van Damme. 1993. "Global Games and Equilibrium Selection." *Econometrica* 61(5): 989-1018.

[2] Chassang, Sylvain, and Gerard Padró i Miquel. 2008. "Conflict and Deterrence under Strategic Risk." NBER Working Paper 13964.

[3] Powell, Robert. 2003. "Nuclear Deterrence Theory, Nuclear Proliferation, and National Missile Defense." *International Security* 27(4): 86-118.