Optimal Political Institutions in the Shadow of Conflict*

Andrea Canidio[†] *and Joan Esteban*[‡]

First version: June 1, 2019. This version: January 8, 2021. Please check here for the latest version.

Abstract

Two players with conflicting interests make investments and then decide whether to trigger a conflict or maintain peace. In case of conflict, these investments determine the players' fighting strength and hence their payoffs. In case of peace, preexisting common political institutions determine the players' payoffs as a function of their investments. We consider the set of political institutions with full information and full commitment and characterize the set of investments compatible with peace. We show that, to maintain peace, the most efficient political institutions may nonetheless distort the players' investments away from the first-best levels. We find conditions under which this distortion is so large, that political institutions capable of maintaining peace do not exist. Therefore, we provide a novel explanation to why rational players may engage in an inefficient conflict, and to why inefficient political institutions exist.

JEL classification: D72, D74, F55.
Keywords: Endogenous political institutions, conflict, international institutions, international agreements.

*We are grateful to Camille Terrier, Vlad Mares, Massimo Morelli, Uwe Sunde, Tim Van Zandt and seminar participants at the INSEAD research symposium, X IBEO Political Economy workshop, First Workshop of CEPR Research and Policy Network on "Preventing Conflict: Policies for Peace", Conference on Mechanism and Institution Design 2020 for their comments and suggestions. Joan Esteban gratefully acknowledges financial support from the AXA Research Fund, the Generalitat de Catalunya, and the Ministry of Economy and Competitiveness Grant number ECO2015-66883-P.

†IMT School for Advanced Studies, Lucca (Italy); andrea.canidio@imtlucca.it.

‡Institut d'Anàlisi Econòmica and Barcelona GSE, Campus UAB, 08193 Bellaterra, Barcelona, Spain; joan.esteban@iae.csic.es.

1 Introduction

One of the most important functions of political institutions is to reconcile competing claims so to avoid conflict. For example, the rules determining political representation and political power indirectly determine how the available resources are allocated among different social, economic, and ethnic groups. An often overlooked observation is that both the surplus to be shared in case of peace and the payoffs in case of conflict are endogenous and depend on prior investments made by those groups. In this paper we show that when those investments are made strategically—that is, anticipating how political institutions will allocate the resulting peace surplus—then there may be no political institution able to prevent the emergence of an inefficient conflict. We therefore propose a novel explanation to why inefficient conflict between rational players may occur.[1]

We model political institutions as an abstract mechanism that allocates the peace surplus between two players as a function of their investments. The key obstacle to peace is that, anticipating the opponent's investment and how political institutions will allocate the peace surplus, one player may find it profitable to invest in improving its conflict payoff and then trigger a conflict. To maintain peace, therefore, the political institutions need to satisfy two endogenous, ex-post participation constraints. The existing literature (which we discuss later) has already noticed that specific political institutions—that is, specific mechanisms to allocate the peace surplus—may fail to satisfy these constraints and therefore lead to conflict. However, when conflict arises under a given mechanism, it is possible that a different mechanism may instead lead to peace. To address this issue, here we characterize the space of all possible political institutions, including those that can credibly commit to destroying part of the peace surplus.

We derive conditions under which there are no political institutions that can achieve the first best. A necessary condition is that the payoff earned by fighting an opponent who expects an efficient peace (and hence invests accordingly) should be larger that the expected payoff earned by fighting an opponent who expects to

[1] In his seminal work, Fearon (1995) lists a number of reasons why rational players may engage in conflict: information asymmetries (see Slantchev, 2010), large indivisibilities, lack of commitment (see Fearon, 1996, Powell, 2006). To these causes Ray (2009) adds the possibility of multilateral bargaining failures. See Jackson and Morelli (2011) for more in depth discussion of the literature on the reasons for conflict.

go to conflict (and hence invests accordingly).[2] Intuitively, if each player expects the other player to choose the first-best investment profile, the benefit of deviating and triggering a conflict may be very large, possibly exceeding the available peace surplus. In this case, to maintain peace, the political institutions may need to distort the players' investment profile away from the first best, so to reduce the incentive to trigger a conflict. Hence, conflict casts a shadow on political institutions, and generates inefficiencies also in case of peace. However, distorting the investment profiles to discourage deviations also reduces the total surplus to be shared in case of peace. We show that, when the underlying conflict is not very destructive, then political institutions that achieve peace may not exist. In this case the unique outcome of the game is an inefficient conflict.

We present a general model and two examples. The first example is a version of Skaperdas (1992)'s "guns and butter" model, in which two players first invest in guns (i.e. weapons) and butter (i.e. productive activities), and then decide whether to trigger a conflict or share the peace surplus via some common political institutions. The peace surplus depends on the total investment in butter, while the investments in guns determine each player's probability of winning the conflict. We show that, to maintain peace, the political institutions may require the players to invest in guns. The "armed peace" discourages the players from triggering a conflict, because each of them anticipates that he will fight an opponent who is armed. Of course, by mandating a positive investment in arms, the political institutions generate an inefficiency. This inefficiency is a function of the destructiveness of the conflict—with more destructive conflict requiring lower investment in guns to prevent deviations and hence generating lower inefficiency. If the destructiveness of the conflict is sufficiently low, it is possible that conflict is the unique equilibrium of the game.[3]

We then modify the example by introducing a second type of productive investment: eggs. Eggs are more costly to produce than butter, but they are more easily destroyed in case of conflict. For this reason, we interpret eggs as state capacity

[2] In technical terms, each players' ex-ante outside option to maintaining peace should be below his ex-post outside option to maintaining peace.

[3] Grossman and Kim (1995) and Skaperdas (1992) consider models in which there are no political institutions, and show that two competing groups may settle for an "armed peace". Here instead the "armed peace" is the constrained efficient outcome, mandated by the optimal political institutions. Note also that in Skaperdas (1992), if the players are symmetric (the case we focus on) there is always conflict, independently of its destructiveness. The existence of common political institutions allows the players, under some conditions, to maintain peace, even if this peace may be "armed".

(which is lost if the state collapses) or human capital (which is lost if people die in a war). Hence, in the first best, the players only invest in butter. To prevent deviations, however, the political institutions may require the players to invest both in eggs and in guns.[4] Interestingly, we show that if the political institutions mandate positive investment in eggs, they also mandate positive investment in guns. Otherwise, a player may deviate by switching 100% of his investment to butter without fear of being attacked—which can be a profitable deviation but, clearly, not an equilibrium of the conflict game.

Finally, two comments on the methodology. We model political institutions as a very abstract mechanism to induce a level of investment and then allocate the resulting peace dividend. We therefore abstract away from a number of frictions and constraints that more realistic political institutions need to face (for example, information frictions, commitment problems, ...). Nonetheless, our results readily extend to those more realistic political institutions: if our abstract mechanism fails to achieve an efficient peace (or to achieve peace at all), then a more realistic mechanism that needs to satisfy additional constrains would also fail to achieve an efficient peace (or to achieve peace at all). Second, for ease of exposition we only consider a finite-time game in which first the players invest, and then there is either conflict or peace. But the model can also be interpreted as a reduced form of an infinitely-repeated game. In this case, the payoff from conflict is the expected present discounted value of deviating one period and then playing conflict in every subsequent period (as in a grim-trigger strategy). The payoff from peace is the expected present discounted value of maintaining peace in every period.[5]

Related literature

The idea that political institutions operate "in the shadow of conflict" is well known in political philosophy, and is central to most theories of the social contract. In particular, in Thomas Hobbes' view, absent political institutions people would live in "the state of nature": the outcome of non-cooperative, violent, rule-free interac-

[4] The fact that government may invest in state capacity to prevent conflict has already been discussed in the literature (see, for example, Besley and Persson, 2009 and Besley and Persson, 2010). What is new here is that *over*investment in state capacity is the outcome under the *optimal* political institutions.

[5] This implicitly characterizes the most efficient equilibrium of the infinitely repeated game but, of course, other equilibria may exist.

tions.[6] Hence the role of political institutions is to provide security and peace. Note that Hobbes' argument readily extends beyond security and peace to all forms of collective action problems, such as for example the provision of public goods (see Taylor, 1987, chapter 1). The possibility of reverting to the state of nature, however, imposes a constraint on the allocations that can be implemented by the political institutions (see Taylor, 1987, chapter 6).

This paper is motivated by the observation that the social surplus to be shared in case of peace and the payoffs in case of conflict (i.e., in the state of nature) depend, at least in part, from prior investments made by the different individual/groups who participate in these political institutions. The endogeneity of these payoffs distinguishes our theory from the existing economic analysis of Hobbes' political philosophy (for example that of Esteban and Sákovics, 2008, Bester and Wärneryd, 2006) and connects us with the literature studying contractual arrangements. In particular we are related to the literature studying contracts with endogenous ex-post outside options.[7] Importantly, here, the ex-post outside option is a conflict, which implies that a player's incentive to deviate (i.e., choose his ex-post outside option) depends on the investments made by *both* players.[8]

As already discussed in the introduction, our paper provides a novel explanation to why rational players may trigger an inefficient conflict: the fact that, anticipating

[6] An even earlier reference to this idea is in Cicero's *Pro Milone*: *Silent enim lēgēs inter arma* (In times of war, the law falls silent).

[7] The most famous model of contracting with endogenous outside option is that of Gibbons and Murphy (1992), in which after signing a labor contract, a worker can take actions that increase his/her outside option. See also Edlin and Reichelstein (1996), Che and Hausch (1999), and Chatterjee and Chiu (2013), in which an agent can make a productive investment that affects both the value of transacting with the other player and the value of transacting with third parties. Kranton and Minehart (2000), Kranton and Minehart (2001) and Elliott (2015) consider a network of buyers and sellers, in which each player can spend resources to link with an additional buyer/seller and therefore increase his bargaining power. Also relevant is Cole, Mailath, and Postlewaite (2001), who study bargaining protocols leading to efficient non-contractible investments prior to matching.

[8] This aspect of the model is similar to Thomas and Worrall (2018), who study an infinitely repeated games in which each player can deviate to an outside option. This outside option is increasing in the action (that is, productive effort) of the other player. They also show that a player's action may be distorted so to discourage the other player from deviating. The main difference is that in Thomas and Worrall (2018) contracts cannot be enforced. The incentive to act cooperatively (i.e., choose a positive action and and then not trigger the outside option) exclusively comes from the infinite horizon of the game. In our framework instead, there is a mechanism (i.e., the political institutions) that determines the players' payoffs as a function of their investments.

how surplus will be shared, the players may make investments aimed at shifting the conflict payoff. However, the benefit of doing so depends on the opponent's investment—and hence on the political institutions. We are therefore related to the literature studying endogenous political institutions. In particular, we show that the optimal political institutions may prevent conflict by mandating an inefficient investment mix. We therefore provide a novel explanation to why inefficient political institutions may exist. This is related to Acemoglu (2003) and Acemoglu (2006), in which elites cannot commit to a given set of transfers, and hence the political outcome may be inefficient. In those papers, however, there is no possibility of conflict. Our argument is therefore related but not identical to that in Acemoglu (2003) and Acemoglu (2006).

Finally, we are related to a number of papers considering specific mechanism to share surplus, and showing that the possibility of investing (i.e., arming) may lead to conflict or to an inefficient peace. For example, in Canidio and Esteban (2018), we consider a very specific family of arbitration procedures and derive the welfare maximizing one. Also related is Meirowitz, Morelli, Ramsay, and Squintani (2019), who compare mediated and unmediated negotiation in a model with pre-negotiation investments, assuming that the mediator behaves like a standard Myerson (1986) mediator. Esteban, Morelli, and Rohner (2015) consider the family of generalized Nash bargaining solutions, and show that an equal surplus-split rule may be welfare decreasing relative to an asymmetric surplus-split rule. Garfinkel, McBride, and Skaperdas (2012) show that when fighting is not sufficiently destructive, arming will be unavoidable within the class of distribution rules they consider. Also related is the model in Anbarci, Skaperdas, and Syropoulos (2002), where each party starts by making wasteful investments in armaments. The paper compares the waste produced by three cooperative bargaining solutions: equal sacrifice, equal benefit, and Kalai-Smorodinski. Both Grossman and Kim (1995) and Skaperdas (1992) assume that there are no political institutions and hence, in the absence of conflict, each player consumers whatever he/she produced. They derive conditions under which there is peace (which will be armed) or conflict.

Here instead we consider the full set of feasible political institutions, which me model as a mechanism to allocate the peace payoff among the two players. We do so by abstracting away from all possible sources of inefficiencies other than the players' ex-post participation constraints

2 General model

We start by presenting a general model that obtains our main results, albeit on a somewhat abstract level. In the next section we consider more specific models and derive additional results.

There are two players, 1 and 2, both having quasilinear utility functions. We interpret the two players as two competing groups. At the beginning of the game, each player $i \in \{1, 2\}$ chooses x_i, which is a vector of L different investments (e.g., transport infrastructures, universities, R&D labs, weapons, military bases, ...) out of a feasible set $X_i \subset \mathbb{R}^L_+$, assumed compact. After simultaneously choosing their investment levels $x_i \in X_i$, each player $i \in \{1, 2\}$ decides whether to trigger a conflict or maintain peace.

Conflict can be unilaterally triggered by either player. In case of conflict, player i's payoff is $w_i(x_1, x_2) : X_1 \times X_2 \to \mathbb{R}_+$.[9] If no player triggers a conflict, then there is peace. In this case, player i's output is $p_i(x_1, x_2) : X_1 \times X_2 \to \mathbb{R}_+$, assumed continuous in all its argument. The common political institutions then implements transfers $T_i(x_1, x_2) : X_1 \times X_2 \to \mathbb{R}$ to each player (positive or negative), which could depend on the players' investments. Each player's peace payoff is therefore $p_i(x_1, x_2) + T_i(x_1, x_2)$. The transfers must satisfy a feasibility constraint, so that $T_1(x_1, x_2) + T_2(x_1, x_2) \leq 0$.[10] Furthermore, they are common knowledge and are fully taken into consideration by the players when performing their initial investments.

Define:

$$\{x_1^*, x_2^*\} \equiv \text{argmax}_{x_1 \in X_1, x_2 \in X_2} \{p_1(x_1, x_2) + p_2(x_1, x_2)\}$$

as the investment levels maximizing aggregate output in case of peace (note that $\{x_1^*, x_2^*\}$ could be a set). Define $x_1^{BR}(x_2)$ and $x_2^{BR}(x_1)$ as the two players' best responses in case of conflict, that is:

$$x_i^{BR}(x_{-i}) \equiv \text{argmax}_{x_i \in X_i} w_i(x_1, x_2).$$

Call x_1^{NE} and x_2^{NE} the Nash equilibrium of the conflict game[11]

$$x_1^{NE} = x_1^{BR}(x_2^{NE}) \qquad x_2^{NE} = x_2^{BR}(x_1^{NE}),$$

[9] Conflict payoffs are therefore assumed positive. This is exclusively for ease of exposition.

[10] Note that, in principle, the optimal political institutions could transfer to the players less than the total output—that is, output could be destroyed. Of course, this will not happen on the equilibrium path, but such threat may be relevant for sustaining the equilibrium investment levels.

[11] Given our assumptions, the Nash equilibrium of the conflict game always exists (see for example

To avoid trivialities, we assume that conflict is inefficient from the ex-ante viewpoint:

$$w_1(x_1^{NE}, x_2^{NE}) + w_2(x_1^{NE}, x_2^{NE}) < p_1(x_1^*, x_2^*) + p_2(x_1^*, x_2^*). \tag{A1}$$

Hence, the first-best level of aggregate output under peace is larger than the sum of payoffs in the Nash equilibrium of the conflict game.

2.1 Optimal political institutions.

We consider here a family of rather extreme political institutions: those that mandate two investment vectors $\bar{x}_1$ and $\bar{x}_2$ (one for each player), and then set transfers so that, if both players comply and maintain peace, the players' payoffs are $\bar{U}_1$ and $\bar{U}_2 \leq p_1(\bar{x}_1, \bar{x}_2) + p_2(\bar{x}_1, \bar{x}_2) - \bar{U}_1$. If a player deviates and peace is maintained, then the political institutions impose the largest possible punishment on the deviating player. Since each player can secure the conflict payoff by triggering conflict, the largest punishment political institutions can credibly impose is to keep the deviating player at his conflict payoff. The non-deviating player receives the rest of the peace surplus $(p_1(x_1, x_2) + p_2(x_1, x_2)) - (w_1(x_1, x_2) + w_2(x_1, x_2))$. If both players deviate and peace is maintained, the Nash bargaining solution is implemented, and each player earns his conflict payoff plus half of the peace surplus.

Because player i peace payoff is $p_i(x_1, x_2) + T_i(x_1, x_2)$, the transfers that implement such political institutions are:

$$T_i(x_1, x_2) =$$
$$\begin{cases} \bar{U}_i - p_i(x_1, x_2) & \text{if } x_i = \bar{x}_i;\ x_{-i} = \bar{x}_{-i} \\ w_i(x_1, x_2) - p_i(x_1, x_2) & \text{if } x_i \neq \bar{x}_i;\ x_{-i} = \bar{x}_{-i} \\ p_{-i}(x_1, x_2) - w_{-i}(x_1, x_2) & \text{if } x_i = \bar{x}_i;\ x_{-i} \neq \bar{x}_{-i} \\ \frac{1}{2}\left(p_{-i}(x_1, x_2) - p_i(x_1, x_2) + w_i(x_1, x_2) - w_{-i}(x_1, x_2)\right) & \text{otherwise,} \end{cases} \tag{1}$$

for $i \in \{1, 2\}$. Note that, in case of deviation by a player, the peace payoff of the non-deviating player may be below his conflict payoff. In this case, the non-

Glicksberg, 1952). Our notation implicitly assumes that this equilibrium is in pure strategy, but our argument holds identical if the equilibrium instead is in mixed strategy—the only difference is that the utility in the equilibrium of the conflict game is now an expectation. If the Nash equilibrium of the conflict game is not unique, we restrict our attention to the Pareto preferred one.

deviating player will trigger a conflict rather than maintaining peace — which is the ex-post efficient outcome. This will also be the case if the peace payoff in case of joint deviation is below the conflict payoff for at least one player. Note also that the above mechanism is ex-post efficient: peace or conflict are reached depending on which outcome is the most efficient ex-post, and in case of peace the entire peace surplus is allocated to the two players.

Note that, for given $\bar{x}_1$, $\bar{x}_2$, $\bar{U}_1$, $\bar{U}_2$, there could be an equilibrium in which both players deviate by investing $\hat{x}_1 \neq \bar{x}_1$ and $\hat{x}_2 \neq \bar{x}_2$, and peace is maintained. Note however that, if such equilibrium exists, then there is also an equilibrium in which $\bar{x}_1 = \hat{x}_1$ and $\bar{x}_2 = \hat{x}_2$, there are no deviations and peace is maintained. That is, an equilibrium with joint deviation is equivalent to an equilibrium without deviations, in which the target investment vectors and the peace utility levels are set appropriately. Similarly, for given $\bar{x}_1$, $\bar{x}_2$, $\bar{U}_1$, $\bar{U}_2$, there could be an equilibrium in which a single player deviates and peace is maintained. Again, this is equivalent to an equilibrium without individual deviations, for some appropriately chosen $\bar{x}_1$, $\bar{x}_2$, $\bar{U}_1$, $\bar{U}_2$. It follows that, when characterizing the equilibrium for every possible $\bar{x}_1$ and $\bar{x}_2$, without loss of generality we can consider exclusively equilibria in which no player has an incentive to make an individual deviation.

Focusing on this family of political institutions has two advantages. First, it is without loss of generality with respect to the equilibrium investment levels: if some political institutions generate the equilibrium investment levels $\tilde{x}_1, \tilde{x}_2$, the same equilibrium investment levels are achieved under the political institutions we consider by setting $\bar{x}_i = \tilde{x}_1$ and $\bar{x}_2 = \tilde{x}_2$. This follows simply from the fact that the benefit from deviating from a given equilibrium investment profile is always (weakly) lower under our political institutions than under any other political institution. Because the family of institutions we consider are ex-post efficient, an interesting implication is that there is no benefit in allowing institutions to implement inefficient outcomes (as a way to punish the players): under our assumptions, institutions that can commit to destroy surplus perform as well as institutions that cannot commit to do so.

Second, we will show that there are conditions under which not even our extreme political institutions can achieve an efficient peace or prevent conflict. This implies that, under the same conditions, *no* political institution can achieve an efficient peace or prevent conflict. For example, it is possible that, because of some information frictions, the payoffs specified by the political institutions cannot depend on certain investments (those that are not observable). Or that only political institutions that are continuous around the target investment levels are feasible (on the ground that the player may make mistakes). All these additional frictions and constraints would add realism to our model, but they are not relevant for our main result. We therefore abstract away from them.

Within this family of political institutions, we are interested in deriving the *optimal* political institutions, that is, the $\bar{x}_1, \bar{x}_2$ maximizing the sum of the players' equilibrium payoffs. Note that these institutions are optimal also among all possible political institutions. This is because, as already discussed, focusing on the above family of political institutions is without loss of generality with respect of the equilibrium investment levels.

2.2 Discussion.

A few aspects of the model deserve to be discussed in some details. We consider political institutions that mandate some investment levels under the threat of a punishment. Therefore, on the equilibrium path, the model is observationally equivalent to a situation in which the investment levels x_1, x_2 are chosen directly by the common political institutions. It is important, however, that each player could deviate to a different investment profile. As we will see, this implies that the mandated investment levels should be incentive compatible.

Also, it is important to clarify that conflict is inefficient from the ex-ante viewpoint: that is, taking into consideration the fact that the equilibrium investments in case of conflict may be different from the first-best ones. However, we do not assume that conflict is inefficient ex-post (that is, for a given investment level). This will depend on the details of the conflict game. For example, it could be that $w_1(x_1, x_2) + w_2(x_1, x_2) = p_1(x_1, x_2) + p_2(x_1, x_2)$, that is, for given investment levels aggregate output is the same under conflict than under peace. In this case, conflict is inefficient whenever each player exercises an externality on the other player, because in case of conflict those investment are chosen non-cooperatively. However, we could also have $w_1(x_1, x_2) + w_2(x_1, x_2) < p_1(x_1, x_2) + p_2(x_1, x_2)$ whenever, if a conflict occurs, the players exert additional fighting effort to capture the aggregate output $p_1(x_1, x_2) + p_2(x_1, x_2)$. Under our assumptions, it is also possible that, for some investment profiles, conflict is preferred to peace. For example, it is possible that after investing heavily in military equipment, it is more efficient to fight a war rather than maintain peace.

Finally, the mechanism considered here is renegotiation proof. As already discussed, the mechanism is ex-post efficient: there is no agreement that can make both players better off relative to what the mechanism assigns them. Note, however, that in case of an individual deviation, the deviating player is indifferent between conflict or maintaining peace—and therefore may threaten to trigger conflict unless the other player makes an additional transfer. There may be an equilibrium where this threat is considered credible, leading to an ex-post reallocation of resources. This equilibrium, however, disappears if the mechanism is such that the deviating player is kept at his conflict payoff plus an arbitrarily small additional transfer, so that he strictly prefers maintaining peace to conflict. For this reason, we focus on the equilibrium in which the threat of triggering conflict when indifferent is not credible. It follows that, even if the players could renegotiate ex-post, they will never want to do that.

2.3 Solution

Suppose player i believes that player $-i$ will follow the prescribed investment level. If player i decides to deviate from $\bar{x}_i$, by (1) it should deviate to $x_i^{BR}(\bar{x}_{-i})$. It follows that there is no profitable deviation from investment levels $\bar{x}_1$ and $\bar{x}_2$ if and only if:

$$\bar{U}_1 \geq w_1(x_1^{BR}(\bar{x}_2), \bar{x}_2) \qquad \text{and} \qquad \bar{U}_2 \geq w_2(\bar{x}_1, x_2^{BR}(\bar{x}_1)).$$

Knowing this, for given $\bar{x}_1$, $\bar{x}_2$ the game is similar to a prisoner's dilemma. The players may be jointly better off by maintaining peace, but if a player expects the opponent to invest the prescribed amount under peace, this player may want to trigger a conflict. Crucially, however, $\bar{x}_1$, $\bar{x}_2$ are chosen endogenously. Hence, both the benefit from cooperation and the incentives to deviate are determined by the investment profiles prescribed by the political institution.

For given $\bar{x}_1$ and $\bar{x}_2$, there exist $\bar{U}_1$, $\bar{U}_2 \leq p_1(\bar{x}_1, \bar{x}_2) + p_2(\bar{x}_1, \bar{x}_2) - \bar{U}_1$ that satisfy both constraints if and only if:

$$p_1(\bar{x}_1, \bar{x}_2) + p_2(\bar{x}_1, \bar{x}_2) \geq w_1(x_1^{BR}(\bar{x}_2), \bar{x}_2) + w_2(\bar{x}_1, x_2^{BR}(\bar{x}_1)). \tag{2}$$

To solve for the optimal political institutions we need to find the $\bar{x}_1$ and $\bar{x}_2$ that maximize $p_1(\bar{x}_1, \bar{x}_2) + p_2(\bar{x}_1, \bar{x}_2)$ subject to (2). If such political institutions exist and generate higher welfare than conflict, then the optimal political institutions are the solution to this problem. Otherwise, the equilibrium of the game is conflict.

We now introduce our main comparative static. For any pair of conflict-payoff functions $w_1(x_1, x_2)$, $w_2(x_1, x_2)$, and any aggregate peace output $p_1(x_1, x_2)+p_2(x_1, x_2)$ satisfying (A1), define

$$\tilde{\alpha} \equiv \frac{p_1(x_1^*, x_2^*) + p_2(x_1^*, x_2^*)}{w_1(x_1^{NE}, x_2^{NE}) + w_2(x_1^{NE}, x_2^{NE})}. \tag{3}$$

Given thism for any $\alpha \in [0, \tilde{\alpha})$ we have:

$$\alpha \left(w_1(x_1^{NE}, x_2^{NE}) + w_2(x_1^{NE}, x_2^{NE}) \right) < p_1(x_1^*, x_2^*) + p_2(x_1^*, x_2^*),$$

which implies that the conflict payoff functions $\alpha w_1(x_1, x_2)$, $\alpha w_2(x_1, x_2)$, and $p_1(x_1, x_2)+p_2(x_1, x_2)$ satisfy (A1).

Therefore, starting from any pair of conflict payoff functions satisfying (A1), we can construct a family of conflict payoff functions satisfying (A1), one for every $\alpha < \tilde{\alpha}$. Furthermore, each conflict payoff function within this family has the same best responses and the same Nash equilibrium, but different payoffs in case of conflict. Hence, varying α allows us to change the destructiveness of the conflict—with lower values corresponding to lower payoffs in case of conflict and hence a more destructive conflict—while leaving everything else (i.e., the best responses of the conflict game and aggregate output in case of peace) constant.

The next proposition states our main result: that if the ex-post outside options are above the ex-ante outside options, then the shadow of conflict may prevent the existence of an efficient peace, and may even prevent the existence of common political institutions.

Proposition 1. *Consider any given* $w_1(x_1, x_2)$*,* $w_2(x_1, x_2)$*, and* $p_1(x_1, x_2)+p_2(x_1, x_2)$ *satisfying* (A1). *If*

$$\begin{aligned} w_1(x_1^{BR}(x_2^*), x_2^*) &\leq w_1(x_1^{NE}, x_2^{NE}) \\ w_2(x_2^{BR}(x_1^*), x_1^*) &\leq w_2(x_1^{NE}, x_2^{NE}) \end{aligned} \tag{4}$$

then political institutions achieving the first best exist for every pair of conflict payoff functions of the form $\alpha w_1(x_1, x_2)$*,* $\alpha w_2(x_1, x_2)$ *for* $\alpha \in [0, \tilde{\alpha})$*.*

If instead

$$\begin{aligned} w_1(x_1^{BR}(x_2^*), x_2^*) &> w_1(x_1^{NE}, x_2^{NE}) \\ w_2(x_2^{BR}(x_1^*), x_1^*) &> w_2(x_1^{NE}, x_2^{NE}) \end{aligned} \tag{5}$$

for every pair of conflict payoff functions $\alpha w_1(x_1, x_2)$, $\alpha w_2(x_1, x_2)$ *we have that*

- *there exists an* $\hat{\alpha} < \tilde{\alpha}$ *such that political institutions achieving the first best exist for* $\alpha < \hat{\alpha}$ *but not for* $\hat{\alpha} < \alpha < \tilde{\alpha}$.
- *if* $\alpha < \tilde{\alpha}$ *but sufficiently close to* $\tilde{\alpha}$, *then no political institutions can prevent conflict, which is therefore the unique outcome.*[12]

To understand the above proposition, note that $w_1(x_1^{NE}, x_2^{NE})$ and $w_2(x_1^{NE}, x_2^{NE})$ are the utilities in case of conflict and hence they are the *ex-ante* outside options: the players' best alternative to setting up common political institutions. Instead $w_1(x_1^{BR}(x_2^*), x_2^*)$ and $w_2(x_1^*, x_2^{BR}(x_1^*))$ are the utility that each player earns by triggering a conflict against an opponent who chose the first-best investment mix. They are the players' *ex-post* outside option to an efficient peace.

The proposition therefore makes clear that, under condition (A1), if the ex-ante outside options are greater than the ex post outside options, then the first best is always achievable. This would be the case if, for example, the payoff in case of conflict is independent of the players' investment. It would also be the case if the first-best level of investment is, for the most part, not appropriable in case of conflict (see Section 3.2 for an example). It corresponds to the "textbook" hold up problem, in which the fact that the ex-post outside option is endogenous is irrelevant, and hence because of full observability and full contractibility the first best is always achieved. When the ex-post outside options are above the ex-ante outside options, instead, whether the first best is achievable depends on how large is the benefit of peace (as measured by α).

If the first best is not achievable, the optimal political institutions will need to distort the investment levels so that (2) is satisfied with equality. This can only be achieved by reducing the RHS of (2). That is, the political institutions will need to distort the investment mix so to make conflict more costly, so to punish each player in case of deviation. This, however, reduces the peace payoff below the first best, and with it the benefit of maintaining peace (the LHS of 2). It is possible that the distortion in the investment mix required to maintain peace is so severe that conflict is preferred to such peace. It is also possible that there is no value of $\bar{x}_1, \bar{x}_2$ that satisfies (2), in which case the only possible outcome is conflict.

[12] Note that the proposition implicitly assumes that $\{x_1^*, x_2^*\}$ is unique. If it is not unique, then the first part of the proposition holds whenever there exists at least one $x_1, x_2 \in \{x_2^*, x_2^*\}$ satisfying (4), and the second part of the proposition holds whenever every $x_1, x_2 \in \{x_2^*, x_2^*\}$ satisfy 5.

The proposition shows that, indeed, if conflict is inefficient but not very costly, then it will be the unique outcome of the game. For intuition, note that for $\alpha \to \tilde{\alpha}$ then total output in the Nash equilibrium of the conflict game is arbitrarily close to the first best. Furthermore, if (5) holds, maintaining peace requires distorting the players investment levels away from the first best. But this means that if α is sufficiently close to $\tilde{\alpha}$, then conflict is preferred to a distorted peace.

Finally, note that whenever neither (4) nor (5) hold, then whether the first best is achievable depends not only on the conflict function $w_i(.,.)$, but also on total output in case of peace $p_1(x_1^*, x_2^*) + p_2(x_1^*, x_2^*)$.

3 Two examples.

To better illustrate the general model, we now present two examples. The first one is a special case of the guns and butter model in Skaperdas (1992), in which there is an investment that is valuable both in peace and in conflict (butter) and one that is valuable only in case of conflict (guns). This example will illustrate how the optimal political institutions may require each player to invest in arms so to discourage the opponent from deviating.[13] This, however, has a cost because it reduces the joint payoff in case of peace. If the destructiveness of the conflict is too low, then preventing deviations this way may be too costly, and the unique outcome of the game is conflict. As the destructiveness of the conflict increases, then fewer resources are necessary to prevent conflict, therefore increasing welfare in case of peace. Hence, this example illustrates that the optimal political institutions may mandate an "armed peace", and also that a more destructive conflict may help sustain peace.

The second example is what we call, a "guns, butter and eggs" model of conflict. In this model, there is a productive investment (eggs) that the players can enjoy only if there is peace (because eggs break in case of conflict). This investment may represent human capital, which is lost in conflict if people are killed. It may also represent goods and services provided by the state, which are lost if common political institutions are destroyed. The result here is that, to prevent conflict, the optimal

[13] As already discussed in the literature review, both Skaperdas (1992) and Grossman and Kim (1995) consider models in which there are no political institutions and show that an armed peace may emerge: the two players may invest in arms but then not engage in conflict. The difference here is that arming is mandated by the optimal political institutions.

political institutions may overinvest in eggs (i.e., publicly provided goods) relative to the first best. Furthermore, we show that if the optimal political institutions mandate positive investment in eggs, they must also mandate positive investment in guns.

3.1 Guns and butter.

The players' investment levels are here $x_i = \{g_i, b_i\}$, where $g_i \geq 0$ are guns and $b_i \geq 0$ is butter. Each player is endowed with a unit of resources, that can be freely allocated to producing either guns or butter, so that $g_1 + b_i = 1$. In case of peace, total surplus to be shared is $b_1 + b_2$. The first-best level of welfare is equal to 2, which is achieved by investing all resources in butter. In case of conflict, instead, player i earns $\alpha(b_1 + b_2)$ with probability $g_i/(g_1 + g_2)$, where $\alpha \geq 0$. If no player invests in guns and a conflict occurs, each player probability of winning is $1/2$. Butter therefore represents investments that are productive both in peace and in case of conflict (but possibly differently so depending on α). Guns instead are non-productive investments that increase the probability of winning a conflict.

Again, the parameter α measures the destructiveness of conflict. For example, the use of guns during a conflict may destroy part of the investment in butter, which implies $\alpha < 1$. If instead $\alpha > 1$, then a given investment in butter generates higher utility in conflict than in peace. We do not think that this last case is particularly realistic,[14] but we will nonetheless consider it in our analysis to illustrate the theoretical possibility that an inefficient conflict is the unique outcome of the game.

Conflict. We start by solving the conflict game. The two best responses are:

$$g_i^{BR}(x_{-i}) = \sqrt{2g_{-i}} - g_{-i}, \qquad b_i^{BR}(x_{-i}) = 1 - g_i^{BR}(x_{-i})$$

The Nash equilibrium is $g_1^{NE} = g_2^{NE} = \frac{1}{2}$, $b_1^{NE} = b_2^{NE} = \frac{1}{2}$. Total output in case of conflict is equal to α. Assumption (A1) holds as long as $\alpha < 2$, which we assume.

Optimal political institutions. To start, note that, here (5) holds and therefore, by Proposition 1, we should expect that for low values of α the first best is achievable,

[14] This is not to say that it is completely unreasonable. For example, it is a known fact that the marginal utility of consumption of some goods increases with the level of stress.

for intermediate values of α the first best is not achievable but may be possible to achieve peace, for high values of α conflict is the unique outcome.

The political institutions set mandatory investment levels $\bar{b}_1 \geq 0, \bar{b}_2 \geq 0, \bar{g}_1 = 1-\bar{b}_1 \geq 0, \bar{g}_2 = 1-\bar{b}_2 \geq 0$ under the threat of keeping a player to his conflict payoff. Equation (2) here is equivalent to:

$$2-(\bar{g}_1+\bar{g}_2) \geq \alpha\left(\sqrt{2}-\sqrt{\bar{g}_1}\right)^2+\alpha\left(\sqrt{2}-\sqrt{\bar{g}_2}\right)^2 \tag{6}$$

The important thing to note is that mandating a given investment in guns decreases each player's incentive to deviate, because each player anticipates that, if he deviates, he will fight against a stronger opponent and the prize in case of victory is smaller. Investing in guns, however, generates a welfare loss and makes maintaining peace less valuable. Simple inspection of the above constraint implies the following proposition.

Proposition 2. *In the guns-and-butter example, the first best is achievable if and only if* $\alpha \leq \frac{1}{2}$.

The above proposition follows by simple inspection of (6). Quite intuitively, at the first-best level of investment, each player can trigger a conflict and capture the entire surplus by investing arbitrarily little in guns. It is possible to prevent both players from deviating only if conflict destroys at least half of the surplus, so that the sum of the utilities from deviating is below the first-best level of welfare.

If, instead $\alpha > \frac{1}{2}$, then conflict is not sufficiently destructive and, as a consequence, the first best is not achievable. Hence the optimal political process will need to impose positive investment in guns, so to make (7) binding.

Proposition 3. *In the guns-and-butter example, whenever* $1/2 < \alpha \leq 1$, *the optimal political institutions maintain peace by imposing*

$$\bar{g}_1 = \bar{g}_2 = \frac{1}{2}\left(\frac{2\alpha-\sqrt{2(1-\alpha)}}{\alpha+1}\right)^2$$

Social welfare is strictly decreasing in α, *and equal to social welfare in case of conflict for* $\alpha = 1$.

Whenever $\alpha \in (1, 2)$ *then it is not possible to satisfy* (6) *and the unique outcome of the game is an inefficient conflict.*

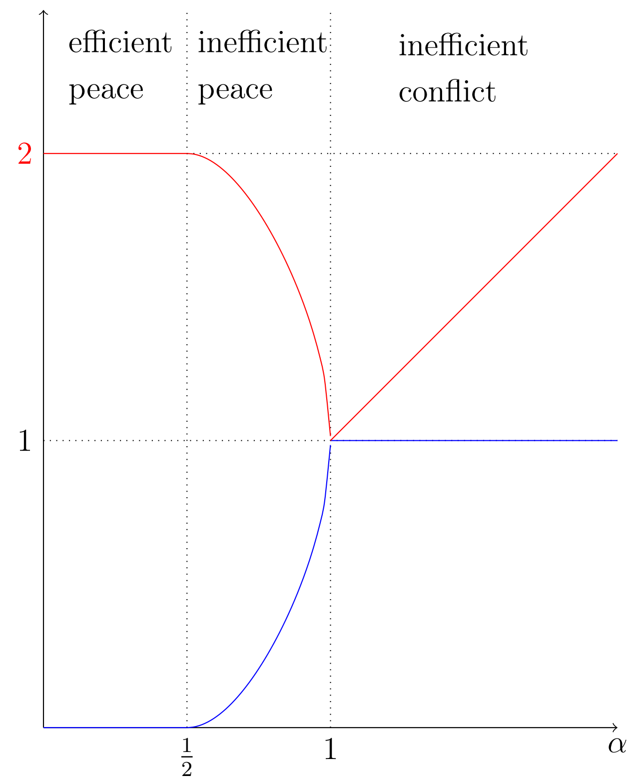

Fig. 1: Equilibrium investment in guns (blue line) and social welfare (red line)

Figure 1 plots total investment in guns and aggregate output in equilibrium. The bottom line is that conflict is avoided by setting up political institutions that require both players to make a positive investment in guns. This investment in guns decreases with the destructiveness of the conflict (as measured by α) because, as α decreases, a smaller investment in guns by a player is required in order to "punish" the other player in case of deviation. It follows that if α is sufficiently small welfare in case of peace achieves the first best level. If instead α is sufficiently large the required punishment is so large that it is not possible to maintain peace, and an inefficient conflict is the only outcome. Contrast this result with Skaperdas (1992), in which there are no political institutions and conflict is the equilibrium of the game for every α.[15] In particular, here, the equilibrium level of welfare is non-monotonic in α: it decreases in α whenever the political institutions can maintain peace, and

[15] This is because the players are identical. In Skaperdas (1992), if players are asymmetric, then conflict may be prevented.

increases in α in case of conflict. In Skaperdas (1992), instead, because conflict is the unique outcome social welfare is always increasing in α.

3.2 Guns, butter and eggs

The general model shows that, in order to maintain peace, the political institutions may distort the players' investment away from the first best level. In the above example this distortion takes the form of requiring players to invest in arms. But, in general, other distortions may emerge. For example, there could be multiple types of productive investments, some more appropriable than others in case of conflict. In this case, it is possible that the political institutions will tilt the investment mix toward the least appropriable productive investment in order to reduce the players' payoff in case of conflict.

To illustrate this possibility, we consider here three types of investments: guns g_i, butter b_i and eggs e_i with $g_i + b_i + e_i = 1$. Eggs are valuable in case of peace, but may be less valuable than butter: total surplus in case of peace is $b_1 + b_2 + \tau(e_1 + e_2)$ for $\tau \geq 0$, where τ is a parameter measuring the marginal rate of technical substitution between butter and eggs in case of peace. In the first best, all resources are invested either in butter (whenever $\tau \leq 1$) or in eggs (whenever $\tau \geq 1$). The resulting first-best social surplus is $2 \max\{1, \tau\}$. In case of conflict, player i earns $\alpha(b_1 + b_2)$ with probability $g_i/(g_1 + g_2)$.

Hence, whereas butter is valuable both in peace an in conflict, eggs are valuable only in peace (because they easily break). For example, butter could represent physical capital while eggs could represent human capital. Eggs could also represent investment in the capacity of the state to provide goods and service, which is lost if the common political institutions are dissolved. Whenever $\tau < 1$ producing eggs is always less efficient than producing butter, and this inefficiency is larger in case of conflict than in case of peace. As we will see, this implies that, to discourage conflict, the optimal political institutions may mandate positive investment in eggs even if $\tau < 1$.

Conflict. The fact that eggs are not valuable in case of conflict implies that the conflict game is a standard "guns and butter" game as in Skaperdas (1992). The two best responses are:

$$g_i^{BR}(x_{-i}) = \sqrt{(2 - e_{-i})g_{-i}} - g_{-i}, \quad b_i^{BR}(x_{-i}) = 1 - g_i^{BR}(x_{-i}), \quad e_i^{BR}(x_{-i}) = 0.$$

The Nash equilibrium is, again, $g_1^{NE} = g_2^{NE} = \frac{1}{2}$, $b_1^{NE} = b_2^{NE} = \frac{1}{2}$, $e_1^{NE} = e_2^{NE} = 0$. Total output in case of conflict is, again, equal to α, and therefore (A1) holds as long as $\alpha < 2\max\{1, \tau\}$.

Optimal political institutions. In case $\tau \geq 1$, condition (4) applies and, by Proposition 1, the first best is achievable. If $\tau < 1$, condition (5) applies and, again by Proposition 1, for low values of α the first best is achievable, for intermediate values of α the first best is not achievable but peace may be possible, for high values of α conflict is the unique outcome. Of course, the difference with the "guns and butter" model presented earlier is that the thresholds determining what case emerges here will depend on τ.

Again, the political institutions set mandatory investment levels $\bar{b}_1, \bar{b}_2, \bar{e}_1, \bar{e}_2, \bar{g}_1, \bar{g}_2$ under the threat of a punishment that cannot exceed a player's conflict payoff. Equation (2) here is equivalent to:

$$2 - (1-\tau)(\bar{e}_1 + \bar{e}_2) - (\bar{g}_1 + \bar{g}_2) \geq \alpha \left(\sqrt{2-\bar{e}_1} - \sqrt{\bar{g}_1}\right)^2 + \alpha \left(\sqrt{2-\bar{e}_2} - \sqrt{\bar{g}_2}\right)^2 \quad (7)$$

Plus two feasibility constraints: $0 \leq \bar{e}_i \leq 1$ and $0 \leq \bar{g}_i \leq 1 - e_i$.

Similarly to mandating a given investment in guns, also mandating a given investment in eggs decreases each player's incentive to deviate. Investing in eggs implies that social surplus is less appropriable by the other player in case of conflict. Also in the case of eggs, however, preventing conflict may come at the cost of reducing the surplus in case of peace.

The next proposition shows that introducing the possibility of investing in eggs expands the range of α for which it is possible to prevent conflict.

Proposition 4. *In the guns-butter-eggs example, if $\alpha \leq 1$ or $\tau \geq 1$ it is always possible to maintain peace (that is, to satisfy 10). If $\alpha \in \left[\frac{\sqrt{5}+1}{2}, 2\right)$ and $\tau < 1$ instead the unique equilibrium is an inefficient conflict. If $\alpha \in \left[1, \frac{\sqrt{5}+1}{2}\right)$ it is possible to maintain peace for $\frac{\alpha}{1-\alpha^2+2\alpha} \leq \tau \leq 1$ but not otherwise.*

Note that the above proposition does not address the question of when the political process will want to maintain peace. That is, it is possible that peace can be maintained but the distortion required is so large that conflict is preferred to peace. We return to this point later (see Corollary 1).

The next proposition provides the full solution for the case $\alpha = 1$ and $\tau < 1$.

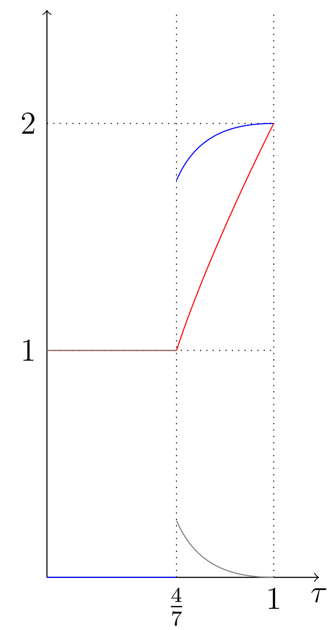

Fig. 2: Equilibrium investment in eggs (blue line), in guns (gray line) and social welfare (red line)

Proposition 5. *Consider the guns-butter-eggs example, and assume $\alpha = 1$, $\tau < 1$. If $\tau < 4/7$ then the solution is again the one derived in Proposition 3: $\bar{g}_1 = \bar{g}_2 = \bar{b}_1 = \bar{b}_2 = 1/2$, $\bar{e}_1 = \bar{e}_2 = 0$, welfare in case of peace is equal to welfare in case of conflict.*

If instead $\tau \geq 4/7$, then

$$\bar{e}_1 = \bar{e}_2 = \frac{3\tau + 2\sqrt{\tau(2\tau - 1)}}{\tau(\tau + 4)} \qquad \bar{g}_1 = \bar{g}_2 = \frac{\tau(\tau + 1) - 2\sqrt{\tau(2\tau - 1)}}{\tau(\tau + 4)} \qquad \bar{b}_1 = \bar{b}_2 = 0$$

welfare in case of peace is strictly greater than welfare in case of conflict, increasing in τ and converging to its first-best level for $\tau \to 1$.

Figure 2 provides a graphical representation of the solution. For low τ the optimal political process will impose positive investment in butter and guns, but no investment in eggs. The solution is therefore the same derived in the previous section. For higher τ instead the investment in eggs will be positive.

Perhaps surprisingly, a positive investment in guns is always required in order to maintain peace, even when the investment in eggs is positive. The reason is that, if all resources are invested in eggs and there is peace, total surplus is 2τ and each player receives τ. A deviating player could instead invest almost all his resources in butter and very little resources in guns, then trigger a conflict and enjoy a utility equal to (approximately) 1. That is, because butter is more productive than eggs, when all resources are invested in eggs a player may deviate not to appropriate the other player's resources, but rather to switch from investing in eggs to investing in butter. As a consequence, if $\tau < 1$ to prevent this deviation some resources will need to be invested in guns.

Finally, the next corollary illustrates that sometimes political institutions could maintain peace, but at the cost of a distortions so large that conflict is preferred to peace.

Corollary 1. *Consider the guns-butter-eggs example with* $\alpha > 1$ *approximately close to 1. Suppose* $\tau \in (1/2, 4/7)$*. Then peace could be maintained but conflict is preferred to peace.*

The fact that peace could be maintained follows directly from Proposition 4. The fact that conflict is preferred to peace follows by continuity to the case $\alpha = 1$ considered in Proposition 5. The only difference is that when $\alpha = 1$ it is possible to maintain peace without investing in eggs, leading to the same social welfare as conflict. If $\tau \in (1/2, 4/7)$, either peace without eggs or conflict are strictly preferred to a peace with eggs. If α is just above 1, instead, it is not possible to maintain peace without eggs. Nonetheless, by continuity welfare in case of conflict is strictly preferred to a peace with eggs.

4 Conclusion

In this paper, we connect Hobbes' political philosophy with modern contract theory. We consider a model in which preexisting political institutions allocate the peace dividend to two players as a function of their investments. Each group can, ex-post, trigger a conflict that dissolves these political institutions. The political institutions are therefore "in the shadow of conflict": the payoff they allocate to the two groups cannot be below what these groups can obtain from conflict. We abstract away from all other forms of frictions and imperfections. Despite this, we find that the first best may not be achievable, in the sense that the optimal political institutions may need to distort the players' investment mix away from the first best in order to obtain peace. It is also possible that the unique outcome of the game is an inefficient conflict. Our contribution is therefore to show that, when players' investments are made anticipating how the political institutions will share the peace dividend, then no political institutions may be able to achieve an efficient peace, or even prevent a conflict. Our results are general in the sense that if the political institutions we consider fail to achieve an efficient peace or to prevent conflict, the same will happen under more realistic political institutions having to satisfy additional constraints.

To better illustrate our results, we consider a guns and butter model à la Skaperdas (1992) and show that the political institutions may require the players to invest in guns. We also consider multiple productive investments, and show that the optimal political institutions may distort the investment mix toward productive investments that are less appropriable in case of conflict. This implies, for example, that to prevent conflict the optimal political institutions may overinvest in goods and services provided by the state (relative to the first best), which are lost if a conflict dissolves the common political institutions.

Appendix

Proof of 1. By evaluating (2) at the first best level of investment, it is immediate to establish that the first best is achievable if and only if:

$$p_1(x_1^*, x_2^*) + p_2(x_1^*, x_2^*) \geq w_1(x_1^{BR}(x_2^*), x_2^*) + w(x_1^*, x_2^{BR}(x_1^*)). \tag{8}$$

By (A1), the above condition is always satisfied whenever 4 holds. This establishes the first part of the proposition.

For the second part, note that 5 implies

$$w_1(x_1^{BR}(x_2^*), x_2^*) + w_2(x_2^{BR}(x_1^*), x_1^*) > w_1(x_1^{NE}, x_2^{NE}) + w_2(x_1^{NE}, x_2^{NE}).$$

The second part of the proposition follows by defining $\hat{\alpha}$ as:

$$\hat{\alpha} \equiv \frac{p_1(x_1^*, x_2^*) + p_2(x_1^*, x_2^*)}{w_1(x_1^{BR}(x_2^*), x_2^*) + w_2(x_2^{BR}(x_1^*), x_1^*)}$$

so that for $\alpha \leq \hat{\alpha}$ then (8) holds, but for $\hat{\alpha} < \alpha < \tilde{\alpha}$ (A1) is satisfied but (8) is violated.

To conclude the proof, assume that $\alpha = \tilde{\alpha}$, so that conflict achieves the first best. If (2) has no solution, then it is not possible to achieve peace. If instead (2) has a solution, by the previous proposition it must be at some $\bar{x}_1, \bar{x}_2$ different from

x_1^*, x_2^*, which implies that peace at $\{\bar{x}_1, \bar{x}_2\} \neq \{x_1^*, x_2^*\}$ is strictly worse than conflict. By continuity, for $\alpha < \tilde{\alpha}$ but approximately close to $\tilde{\alpha}$, either (2) has no solution (in which case the only possible outcome is conflict) or (2) has a solution which is however strictly worse than conflict. □

Proof of Proposition 3. Call G the total investment in guns, with β the fraction of G invested by player 1, so that $\bar{g}_1 = \beta G$. Constraint (6) becomes

$$G(1+\alpha) + 2(2\alpha - 1) \leq 2\alpha\sqrt{G}(\sqrt{2\beta} + \sqrt{2(1-\beta)}) \tag{9}$$

If the LHS of the above inequality crosses its RHS, it will actually cross twice. The smallest G that satisfies (10) is the smallest of such intercepts, where the LHS of (10) crosses its RHS from below. This G is minimized whenever the RHS of (10) is maximized, which happens at $\beta = 1/2$. At this β (10) becomes

$$G(1+\alpha) + 2(2\alpha - 1) \leq +4\alpha\sqrt{G}$$

with solution

$$G^{NE} \equiv \left(\frac{2\alpha - \sqrt{2(1-\alpha)}}{\alpha + 1}\right)^2$$

If $\alpha \leq 1$, the above solution always exists. It is also easy to check that social welfare in case of peace is strictly greater than social welfare in case of conflict for $\alpha < 1$ and is equal to social welfare in case of peace for $\alpha = 1$. Social welfare in case of peace is also strictly decreasing (for a numerical solution, see Figure 1).

If instead $\alpha > 1$, then G^{NE} does not exist. It is not possible to satisfy (10) and hence conflict is the only outcome.

□

Proof of Proposition 4. Call G the total investment in guns, with β the fraction of G invested by player 1, so that $\bar{g}_1 = \beta G$. Call E the total expenditure in eggs, with γ the fraction invested by player 1, so that $\bar{e}_1 = \gamma E$. Constraint (7) becomes

$$G(1+\alpha) + 2(2\alpha - 1) \leq (\tau + \alpha - 1)E + 2\alpha\sqrt{G}(\sqrt{(2-\gamma E)\beta} + \sqrt{(2-(1-\gamma)E)(1-\beta)}) \tag{10}$$

We fix E and look for the smallest G that satisfies the above constraint for some β and γ.

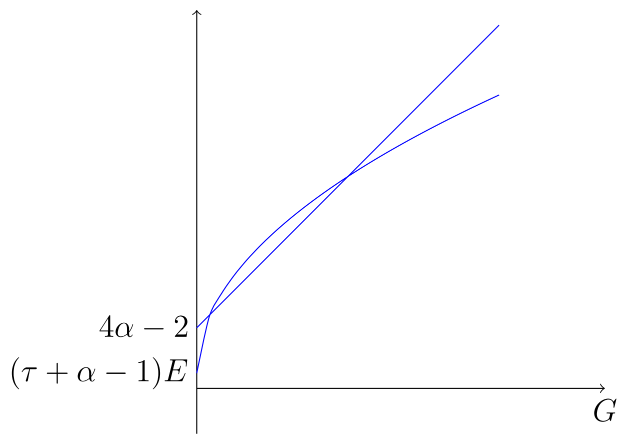

Fig. 3: LHS and RHS of 10, case 2.

We distinguish between to cases. Case 1 is:

$$2(2\alpha - 1) \leq (\tau + \alpha - 1)E,$$

In this case (10) holds at $G = 0$, which is therefore the welfare maximizing G for given E, independently from γ and β. Note that if $\alpha \leq 1/2$ the above inequality holds at $E = 0$, which implies that the first best is achievable. If instead $\alpha > 1/2$ (which is what we assume here) then if $\tau + \alpha \leq 1$ the above inequality never holds and hence we are never in this case. If $\tau + \alpha > 1$ the above inequality holds for $E \geq \frac{2(1-2\alpha)}{\alpha+\tau-1}$. Note that because E is chosen optimally, then this case can emerge because $E = \frac{2(1-2\alpha)}{\alpha+\tau-1}$, which is also a subcase of case 2 (below).

Case 2 is:

$$2(2\alpha - 1) \geq (\tau + \alpha - 1)E.$$

In this case the LHS of (10) crosses its RHS twice (see Figure 3). The smallest G that satisfies (10) is the smallest of such intercepts, where the LHS of (10) crosses its RHS from below. This G is minimized whenever the RHS of (10) is maximized. For given γ, the RHS is maximized at $\beta = (2 - \gamma E)/(4 - E)$. By plugging this value of β into the RHS of 10, we see that the γ drops out. There are therefore multiple possible combinations of γ and β that maximize the RHS of (10). Among these solutions, the one at which the feasibility constraint is more likely to hold is $\beta = \gamma = 1/2$. At those γ and β (10) becomes

$$G(1 + \alpha) + 2(2\alpha - 1) \leq (\tau + \alpha - 1)E + 2\alpha\sqrt{G(4 - E)}$$

with solution

$$G \geq G(E) \equiv \left(\frac{\alpha\sqrt{4-E} - \sqrt{E(\tau(1+\alpha)-1)+2(1-\alpha)}}{\alpha+1}\right)^2$$
$$= \frac{E((\alpha+1)\tau - \alpha^2 - 1) + 2(2\alpha^2 - \alpha + 1) - 2\alpha\sqrt{A(E)}}{(\alpha+1)^2}$$

where

$$A(E) \equiv (4-E)\left[E(\tau(1+\alpha)-1) + 2(1-\alpha)\right]$$

The key observation is that $G(E)$ may not exist. When this is the case, there is no political process that satisfies the no-deviation constraint. The existence of $G(E)$ for some feasible E depends on cases:

- If $\alpha \leq 1$, then $G(E)$ exists for some feasible E. To see this, just consider $E = 0$, so that $G(0) = \left(\frac{2\alpha - \sqrt{2(1-\alpha)}}{\alpha+1}\right)^2$ which is feasible because simple algebra shows that $G(0) < 2$.

- If $\alpha > 1$ and $\tau \leq \frac{1}{1+\alpha}$ then $G(E)$ never exists. If instead $\tau > \frac{1}{1+\alpha}$ then $G(E)$ exists for E sufficiently large. The E such that $G(E)$ exists may, however, not be feasible. To see this, consider the smallest E such that $G(E)$ exists (that is, such that $A(E) = 0$): $E = 2\frac{\alpha-1}{\tau(\alpha+1)-1}$. At this E the feasibility constraint holds if:

  $$E + G(E) = 2\frac{\alpha-1}{\tau(\alpha+1)-1}\left(1 + \frac{(\alpha+1)\tau - \alpha^2 - 1}{(1+\alpha)^2}\right) + \frac{2(2\alpha^2 - \alpha + 1)}{(1+\alpha)^2} \leq 2$$

  or

  $$\tau \geq \bar{\tau} \equiv \frac{\alpha}{1 - \alpha^2 + 2\alpha}$$

  To conclude, note that $\bar{\tau} > \frac{1}{1+\alpha}$ whenever $\alpha > 1$. Hence, whenever $\alpha > 1$, $\tau \geq \bar{\tau}$ guarantees the existence of $G(E)$ at some E. Also, because τ must be below 1, $\tau \geq \bar{\tau}$ never holds if $\alpha \geq \frac{\sqrt{5}+1}{2}$.

□

Proof of Proposition 5. In the proof of Proposition 4 we derived $G(E)$, which, if $\alpha = 1$, becomes

$$G(E) = \frac{2 - E(1-\tau) - \sqrt{(4-E)E(2\tau-1)}}{2}$$

If $\tau < 1/2$, then the only solution is $E = 0$, $G = 1$, which is the same solution derived in the model without eggs. If instead $\tau \geq 1/2$ then $G(E)$ exists for all feasible values of E.

The value of E is chosen to to minimize $G(E) + (1-\tau)E$, which here becomes

$$1 + (1-\tau)E - \frac{\sqrt{(4-E)E(2\tau-1)}}{2}$$

There are three possible solutions: a corner solution at $E = 0$, a corner solution at $E + G(E) = 2$, and an interior solution.

Taking first order conditions we get

$$E = 2 + \frac{1-\tau}{\sqrt{4\tau^2 - 6\tau + 3}}$$

which is, however, not feasible.

At the corner solution $E = 0$ we are back at the case without eggs, and social welfare is 1.

At the other corner solution we have $E + G(E) = 2$ or

$$E = \frac{6\tau + 4\sqrt{\tau(2\tau-1)}}{\tau(\tau+4)} \quad G = \frac{2\tau(\tau+1) - 4\sqrt{\tau(2\tau-1)}}{\tau(\tau+4)}$$

and no investment in butter. Social welfare is

$$\tau\frac{6\tau + 4\sqrt{\tau(2\tau-1)}}{\tau(\tau+4)}$$

which is greater than 1 only if $\tau > 4/7$.

□

References

Acemoglu, D. (2003). Why not a political coase theorem? social conflict, commitment, and politics. *Journal of comparative economics 31*(4), 620–652.

Acemoglu, D. (2006). A simple model of inefficient institutions. *Scandinavian Journal of Economics 108*(4), 515–546.

Anbarci, N., S. Skaperdas, and C. Syropoulos (2002). Comparing bargaining solutions in the shadow of conflict: how norms against threats can have real effects. *Journal of Economic Theory 106*(1), 1–16.

Besley, T. and T. Persson (2009). The origins of state capacity: Property rights, taxation, and politics. *American economic review 99*(4), 1218–44.

Besley, T. and T. Persson (2010). State capacity, conflict, and development. *Econometrica 78*(1), 1–34.

Bester, H. and K. Wärneryd (2006). Conflict and the social contract. *Scandinavian Journal of Economics 108*(2), 231–249.

Canidio, A. and J. Esteban (2018). Benevolent arbitration in the shadow of conflict. *working paper*.

Chatterjee, K. and Y. S. Chiu (2013). Bargaining, competition and efficient investment. In *Bargaining in the Shadow of the Market: Selected Papers on Bilateral and Multilateral Bargaining*, pp. 79–95. World Scientific.

Che, Y.-K. and D. B. Hausch (1999). Cooperative investments and the value of contracting. *American Economic Review 89*(1), 125–147.

Cole, H. L., G. J. Mailath, and A. Postlewaite (2001). Efficient non-contractible investments in finite economies. *Advances in Theoretical Economics 1*(1).

Edlin, A. S. and S. Reichelstein (1996). Holdups, standard breach remedies, and optimal investment. *The American Economic Review 86*(3), 478–501.

Elliott, M. (2015). Inefficiencies in networked markets. *American Economic Journal: Microeconomics 7*(4), 43–82.

Esteban, J., M. Morelli, and D. Rohner (2015). Strategic mass killings. *Journal of Political Economy 123*(5), 1087–1132.

Esteban, J. and J. Sákovics (2008). A theory of agreements in the shadow of conflict: the genesis of bargaining power. *Theory and Decision 65*(3), 227–252.

Fearon, J. D. (1995). Rationalist explanations for war. *International organization 49*(3), 379–414.

Fearon, J. D. (1996). Bargaining over objects that influence future bargaining power. *Working paper*.

Garfinkel, M. R., M. McBride, and S. Skaperdas (2012). Governance and norms as determinants of arming. *Revue d'économie politique 122*(2), 197–212.

Gibbons, R. and K. J. Murphy (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of political Economy 100*(3), 468–505.

Glicksberg, I. L. (1952). A further generalization of the kakutani fixed point theorem, with application to nash equilibrium points. *Proceedings of the American Mathematical Society 3*(1), 170–174.

Grossman, H. I. and M. Kim (1995). Swords or plowshares? a theory of the security of claims to property. *Journal of Political Economy 103*(6), 1275–1288.

Jackson, M. O. and M. Morelli (2011). The reasons for wars: an updated survey. In C. J. Coyne and R. L. Mathers (Eds.), *The handbook on the political economy of war*, pp. 34–57. Edward Elgar Publishing.

Kranton, R. E. and D. F. Minehart (2000). Networks versus vertical integration. *The Rand journal of economics*, 570–601.

Kranton, R. E. and D. F. Minehart (2001). A theory of buyer-seller networks. *American economic review 91*(3), 485–508.

Meirowitz, A., M. Morelli, K. W. Ramsay, and F. Squintani (2019). Dispute resolution institutions and strategic militarization. *Journal of Political Economy 127*(1), 378–418.

Myerson, R. B. (1986). Multistage games with communication. *Econometrica: Journal of the Econometric Society*, 323–358.

Powell, R. (2006). War as a commitment problem. *International organization 60*(1), 169–203.

Ray, D. (2009). Costly conflict under complete information. *Manuscript, Dept. Econ., New York Univ.*

Skaperdas, S. (1992). Cooperation, conflict, and power in the absence of property rights. *The American Economic Review*, 720–739.

Slantchev, B. L. (2010). Feigning weakness. *International Organization 64* (3), 357–388.

Taylor, M. (1987). *The Possibility of Cooperation.* Cambridge University Press.

Thomas, J. P. and T. Worrall (2018). Dynamic relational contracts under complete information. *Journal of Economic Theory 175*, 624 – 651.