Article

A Game-Free Microfoundation of Mutual Optimism

Marco Serena

Max Planck Institute for Tax Law and Public Finance, Department of Public Economics, Marstallplatz 1, D-80539 München, Germany; marco.serena@tax.mpg.de

Received: 11 July 2019; Accepted: 25 September 2019; Published: 27 September 2019

Abstract: One of the most widely accepted explanations for why wars occur despite its Pareto-suboptimality is mutual optimism: if both sides expect to gain a lot by fighting, war becomes inevitable. The literature on mutual optimism typically assumes mutually optimistic beliefs and shows that, under such an assumption, war may occur despite its Pareto-suboptimality. In a war–peace model, we show that, if players neglect the correlation between other players' actions and their types—a well-established concept in economics—then players' expected payoffs from war increase relative to conventional informational sophistication predictions, hence providing a microfoundation of mutual optimism.

Keywords: mutual optimism; incentives to go to war; information; correlation neglect

1. Introduction

If rational country leaders have mutually consistent beliefs about the outcome of a costly war, then a bargain in which Pareto improves upon war must be reachable. The lack of mutually consistent beliefs is an often mentioned rationalist explanation for war. As reported by Slantchev and

If both sides expect to gain a lot by fighting—perhaps because both expect to win with near certainty at an acceptably low cost—then [...] war becomes the inevitable outcome. This argument is now generally known as the mutual optimism explanation of war and is among the most widely accepted explanations of why war occurs.

Due to mutual optimism, both players may expect to be better off going to war even if a war would shrink the players’ aggregate payoff. Pioneers of this idea are the seminal works by Wittman (1979) [2] and Blainey (1988) [3], and a number of scholars have contributed to the idea (e.g., Morrow, 1989 [4]; Fearon, 1995 [5]; Werner, 1998 [6]; Wagner, 2000 [7]; Slantchev and Tarar, 2011) [1]. This strand of the literature typically assumes mutual optimism in war–peace models and concludes that mutual optimism enhances the incentives to go to war: “it could be that two states each are optimistic and are convinced that they will benefit from a war. In these cases war can erupt, as long as the inconsistency of beliefs is large enough to compensate for the cost of war” (Jackson and Morelli, 2011) [8]. In the present paper, we provide a microfoundation for mutual optimism: namely, we show that mutual optimism arises when players correctly predict the distribution of other players' actions and types but—in contrast to informational sophistication—draw no inference about the correlation between the two.1

Such informational naivety of players has a long and prosperous history in economics: pioneered, among others, by Kagel and Levin (1986) [9] and Holt and Sherman (1994) [10], it has been extended

1 In line with orthodox economic terminology, type refers to payoff-relevant parameters.
along several directions. For instance, informational naivety is a special case of cursed equilibria (Eyster and Rabin, 2005) [11], analogy-based expectation equilibria (Jehiel, 2005) [12], and self-confirming equilibria (Fudenberg and Levine, 1993) [13].

Hence, the informational naivety of the present paper reaches out to all those concepts. Since, to the best of our knowledge, there is no unique undisputed terminology for the correlation neglect between other players’ actions and types, we choose the neutral though novel term informational naivety.

We incorporate informational naivety of players into a war–peace model. The main result of the paper is to show that informational naivety increases players’ expected payoffs from war relative to informational sophistication predictions, thus microfounding mutual optimism.

The following stylized version of our war–peace model illustrates the role of informational naivety. Each of the two players has resources (types) independently drawn from a prior distribution. A player privately knows their own resources but ignores the amount of resources of the other. Under peace, players consume their own resources, but if a war breaks out, the winner obtains the total resources and the loser obtains 0. One common modeling approach is that a player, upon observing their own resources, chooses their level of military action, which increases their probability of winning a prospective war. However, since many war–peace models predict a positive equilibrium relation between a player’s resources and their military action, one could sidestep the explicit modeling of military actions and directly assume that a player’s probability of victory increases in their resources. We make this monotonicity assumption in the main part of the present paper (until Section 6), since for our purposes, we are not interested in players’ choice of military actions but in the characterization of players’ expected payoffs from war with and without informational naivety. This monotonicity assumption, named (MonProb) in the model (see Section 3), is both common in the literature and empirically sound.

Superimposing to the present model an endogenous military action which fulfills the abovementioned monotonicities would thus not affect our result. In fact, Section 7 endogenizes efforts, confirms the result of the previous sections, and derives comparative statics.

Therefore, as we do not model military actions explicitly, we implement informational naivety as neglecting the correlation between types and the result of strategic interactions yielding a certain probability of victory. In particular, an informationally naïve player correctly predicts the distribution of possible levels of resources of their rival and of the possible probabilities of victory, but in contrast to informational sophistication, they draws no inference about the correlation between the two: in other words, an informationally naïve player fails to infer the mapping from each possible level of resource of their rival to their corresponding probability of victory. We show that this failure systematically increases players’ expected payoffs from war; that is, informational naivety microfounds mutual optimism.

Section 2 analyses the most stylized example of our model capable of capturing our result and explains its simple intuition. Sections 3 and 4 provide a formal generalization of the example in Section 2. Section 5 discusses four further extensions. Section 6 discusses the model’s lack of choice

2 In a cursed equilibrium, players draw partial inference about the correlation between other players’ actions and their types. Informational naivety is the fully cursed equilibrium, whereby no inference is drawn. Informational naivety is also a special case of an analogy-based expectation equilibria, where players’ analogy partitions coincide with their own information partitions (see Eyster and Rabin, 2010, p. 1634 [14]; Jehiel and Koessler, 2007, p. 539 [15]; and Ettinger and Jehiel, 2010 [16], footnote 7). Finally, informational naivety is a special case of self-confirming equilibria, where the cursed players observe only the aggregate play of the opponents and neither the state nor the opponent’s type (Fudenberg, 2006) [17]. The three concepts differ in how they convexify informational naivety and informational sophistication. Since we focus only on these two extremes, we do not need to take a stand among the three concepts.

3 In the general model of Section 3, we will only adopt the weaker assumption, (MonSpoils), that the spoils of war increase in the rival’s resources.

4 For instance, Jackson and Morelli (2007) [18] assume that the probability of winning the war satisfies (MonProb) without explicitly modeling the arming phase. Bueno de Mesquita (1981, p. 102) [19] also assumes that wealth translates into military capability. Hörner et al.’s (2015) [20] workhorse model has two types, h and l, and probabilities satisfy $p_{h, l} > 1/2 = p_{l, h}$, consistent with (MonProb). Taking types as sunk military investments, Meirovitz and Sartori (2008) [21] assume (MonProb). Furthermore, (MonProb) typically arises in conflict models with resource constraints, such as Tullock contests and Colonel Blotto games. For empirical evidence supporting (MonProb), see, for instance, footnote 15 in Jackson and Morelli (2007) [18] and the references therein.
variables for players. Conversely, Section 7 analyses a simple model where players choose efforts and derives comparative statics. Section 8 discusses our result.

2. A Stylized Example and the Intuition

In order to spell out the intuition, we present the most stylized nontrivial example of the model of Section 3.

Each of the two players privately knows their own resources, which are either high \((R > 0)\) or low \((0)\) with equal probability. Under peace, each player consumes their own resources. Under war, the winner obtains the total amount of resources minus the costs of war \(c \geq 0\) and the loser obtains 0. When a war is between players with equal resources, each has a 1/2 probability of winning. When a war is between players with unequal resources, the player with resources \(R\) has a probability of winning equal to \(p \in (1/2, 1]\). This specification is consistent with \((\text{MonProb})\): win probabilities increase in one’s own type.

Under informational sophistication (IS), a player with resource \(R\) expects to be better off under war than peace when

\[
\pi_{\text{IS}}^R \equiv \frac{1}{2}\left[\frac{1}{2}(2R - c)\right] + \frac{1}{2}[p(R - c)] \geq R. \tag{1}
\]

With probability 1/2, a player with resources \(R\) is up against a rival also with resources \(R\), so that they have a 1/2 probability of winning, \(2R - c\) (the total resources minus the costs of war), while with probability 1/2, a player with resources \(R\) is up against a player with 0 resources, so that they have probability \(p\) of winning \(R - c\) (the total resources minus the costs of war).

Under informational naivety (IN), a player with resources \(R\) expects to be better off under war than peace when

\[
\pi_{\text{IN}}^R \equiv \left(\frac{1}{2} + \frac{1}{2}p\right)\left(\frac{1}{2}2R + \frac{1}{2}R - c\right) \geq R. \tag{2}
\]

Despite an informationally naive player correctly perceiving their average probability of victory (first bracket of Equation (2)) and their average spoils of war (second bracket of Equation (2)), they draw no inference about the correlation between the two.

An IN player differs from an IS player in that they fail to understand that high spoils of war \((2R)\) bring along the bad news that their probability of victory is only 1/2 rather than \(p \in (1/2, 1]\) and that low spoils of war \((R)\) bring along the good news that their probability of victory is \(p \in (1/2, 1]\). These two forces go in the same direction: an IN player with resources \(R\) overestimates their expected payoff from war relative to an IS player. Formally, \(\pi_{\text{IN}}^R > \pi_{\text{IS}}^R\) for every parameter triple \((R, c, p)\). The same holds for an IN player with resources 0: \(\pi_{\text{IN}}^0 > \pi_{\text{IS}}^0\) for every parameter triple \((R, c, p)\).

In words, informational naivety increases a player’s expected payoff from war relative to informational sophistication predictions.

Despite the main point of the paper being that \(\pi_{\text{IN}}^R > \pi_{\text{IS}}^R\) and \(\pi_{\text{IN}}^0 > \pi_{\text{IS}}^0\), it may still be of interest to analyse when players are better off under war than under peace. Trivial algebra gives the threshold for \(c\) below which a player is better off under war:

\[
\pi_{\text{CGT}}^0 = \frac{1}{2}\left[\frac{1}{2}(-c)\right] + \frac{1}{2}[(1 - p)(R - c)] \geq 0
\]

\[
\pi_{\text{IN}}^0 = \left(\frac{1}{2} + \frac{1}{2}(1 - p)\right)\left(\frac{1}{2}R - c\right) \geq 0
\]

---

5 The normalization of low resources to 0 is qualitatively innocuous.
6 The result would carry over if the costs of war \(c \geq 0\) depend on players’ resources or are paid by both players rather than only by the winner. Similarly, the normalization of the loser’s payoff to 0 is without loss of generality, and for simplicity, the normalization is maintained throughout the paper.
7 Peace payoffs depend on resources but not on probabilities of victory and are thus unaffected by informational naivety.
8 The conditions corresponding to Equations (1) and (2) for a player with 0 resources read as follows:

\[
\pi_{\text{CGT}}^0 \equiv \frac{1}{2}\left[\frac{1}{2}(-c)\right] + \frac{1}{2}[(1 - p)(R - c)] \geq 0
\]

\[
\pi_{\text{IN}}^0 = \left(\frac{1}{2} + \frac{1}{2}(1 - p)\right)\left(\frac{1}{2}R - c\right) \geq 0
\]
From Table 1, we can draw a number of conclusions. If the costs of war are overwhelmingly high, peace is the unique outcome of virtually all standard war-declaration models superimposed on the above payoff structure, since both players are strictly better off under peace. Similarly, overwhelmingly low costs of war yield a war to break out. The predictions are more interesting when costs are intermediate, where war is sustained according to resources and informational naivety (IN). Since \( c^{IS}_0 < c^{IS}_1 \), there is an intermediate range of costs, \( c \in [c^{IS}_0, c^{IS}_1] \), where a 0-resource player prefers war than peace under IN but not under IS. Similarly, since \( c^{IN}_0 < c^{IN}_1 \), an R-resource player prefers war than peace under IN but not under IS. Notice that, while for 0-resource players, \( 0 < c^{IN}_0 \), for R-resource players, \( c^{IN}_R \leq 0 \) and \( c^{IN}_R \geq 0 \iff p \geq 5/6 \). Hence, while the intermediate range of costs always exists for 0-resource players, it may not exist for R-resource players.

<table>
<thead>
<tr>
<th>Table 1. Threshold of ( c ) below which a player expects to be better off under war than peace.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Better off under war if</strong></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Informational Sophistication</td>
</tr>
<tr>
<td>Informational Naivété</td>
</tr>
</tbody>
</table>

The stylized example has been conceived under (i) distribution of types, which is uniform, binary, and symmetric across players; (ii) spoils of war, which equal the sum of types minus the fixed cost of war; and (iii) probabilities of victory taking values \( 1/2, p, \) or \( 1 - p \). However, the result of the stylized example is no coincidence, and Sections 3 and 4 generalize it to (i) possibly nonuniform \( n \)-type asymmetric distribution of types, (ii) spoils of war decreasing in rival’s type, and (iii) probability of victory increasing in one’s own type.

3. The Model

Each of the two risk-neutral players privately observes their own type; in particular, the first (or second) player’s type is drawn from a distribution, which assigns probability \( r_i \in [0,1] \) \( q_i \in [0,1] \) to each type \( \theta_i \) with \( i \in \Theta \equiv \{1, \ldots, n\} \) and \( \sum_{i \in \Theta} r_i = 1 \) \( \sum_{i \in \Theta} q_i = 1 \). This specification allows players’ type distributions to possibly differ in probabilities and supports (e.g., setting \( r_1 > 0 = q_1 \) makes type \( \theta_1 \) possible only for the first player). Without loss of generality, assume that \( \theta_1 < \theta_2 < \ldots < \theta_n \), and we refer to type \( \theta_n \) as the strongest (wealthiest) type. We compare two alternative settings: peace, \( \mathcal{P} \), and war, \( \mathcal{W} \).

Under \( \mathcal{P} \), we denote player \( i \)’s payoff under peace by \( \pi^P_i (\theta_i, \theta_j) \). If types are resource levels (e.g., territory, GDP, and technology), as in the example in Section 2, \( \pi^P_i (\theta_i, \theta_j) = \theta_i \); that is, under peace, a player consumes their own resources. We do not need to specify how \( \pi^P_i \) depends on types: thus, for instance, we allow for spillovers across types as well as different returns to types.

Under \( \mathcal{W} \), players engage in a war. If player \( i \) loses, their payoff is 0. If player \( i \) wins, their payoff is \( \pi^W_i (\theta_i, \theta_j) \equiv f (\theta_i, \theta_j) > 0 \). As discussed, one’s spoils of war increase in the rival’s type. Formally,

\[
(MonSpoils) : f (\theta_i, \theta_j) \text{ strictly increases in } \theta_j \text{ for each } \theta_i.
\]

The interpretation of \( (MonSpoils) \) clearly depends on the interpretation of type. If types are resource levels, as in our leading example, players engage in a resource war, and \( (MonSpoils) \) says that the richer the defeated rival, the greater the spoils of war. A special case is that the winner obtains the total resources, possibly with destructiveness parameters \( d_1, d_2 \in [0,1] \) and/or a cost of

\[^9\] Notice that, while for 0-resource players, \( 0 < c^{IS}_0 < c^{IS}_1 \), for R-resource players, \( c^{IS}_R \leq 0 \) and \( c^{IS}_R \geq 0 \iff p \geq 5/6 \). Hence, while the intermediate range of costs always exists for 0-resource players, it may not exist for R-resource players.
war \( c \geq 0 \)\(^{10}\) that is, \( f (\theta_i, \theta_j) = d_1 \theta_i + d_2 \theta_j - c \).\(^{11}\) In Section 2, we discussed \( d_1 = d_2 = 1 \). Alternatively, one could interpret types as military proficiencies, troop qualities, or political resolves: the benefit of war obtained by a player who defeats a stronger or more resolute rival than themself is in terms of glory or reputational gains and, thus, depends positively on the rival’s type and negatively on one’s own type: for instance, \( f (\theta_i, \theta_j) = \theta_i - \theta_j \) or \( f (\theta_i, \theta_j) = \theta_i/\theta_j \).

Furthermore, we assume that, in case of war \( W \), a player’s probability of victory increases in their own type. Formally, if we denote by \( p_i (\theta_i, \theta_j) \) the probability of winning the war of a player of type \( \theta_i \) up against a player of type \( \theta_j \), then we impose the following monotonicity assumption:\(^{12}\)

\[
(MonProb) : p_i (\theta_i, \theta_j) \text{ strictly increases in } \theta_j \text{ for each } \theta_i.
\]

Exactly one of the two players wins the war: if players are of types \( \theta_i \) and \( \theta_j \), then \( p_i (\theta_i, \theta_j) + p_j (\theta_j, \theta_i) = 1 \). An immediate consequence of this fact, together with \((MonProb)\), is that a player’s probability of victory strictly decreases in their rival’s type: \( p_i (\theta_i, \theta_j) \) strictly decreases in \( \theta_j \).

If types are resource levels, as in our leading example, \((MonProb)\) echoes, for instance, Jackson and Morelli (2007)\(^ {18}\) and the papers discussed in footnote 4. An examples of conflict technology consistent with \((MonProb)\) is the Tullock success function, where \( p_i (\theta_i, \theta_j) = \theta_i / (\theta_i + \theta_j) \). However, our main result, much as the majority of Jackson and Morelli’s (2007)\(^ {18}\) results, do not depend on which specific conflict technology is chosen as long as \((MonProb)\) is fulfilled.

4. The Main Result

As \( \pi^W_i (\theta_i, \theta_j) \) depends on types but not on probabilities of victory, a player’s expected payoff under peace is identical under IN and IS. Therefore, we exclusively focus on whether and how players’ expected payoffs from war vary between IN and IS.

Proposition 1. Assume \((MonProb)\) and \((MonSpoils)\). A player’s expected payoff from war is strictly greater when they are informationally naïve than under informational sophistication.

Proof of Proposition 1. We want to show that \( E_{IN} [\pi^W_i] - E_{IS} [\pi^W_i] > 0 \). Let the first player’s type be a generic \( \theta_i \): their probability of winning a war is \( p_i (\theta_i, \theta_j) \) when up against a \( \theta_j \) rival. Indeces \( i, j, \) and \( k \) will be used interchangeably throughout the algebraic steps below.

The difference of expected payoffs from war for the first player of generic type \( \theta_i \) under IN and IS is positive if

\[
E_{IN} [\pi^W_i] - E_{IS} [\pi^W_i] > 0,
\]

\[
\sum_{j=1}^{n} q_j p_i (\theta_j, \theta_j) \left( \sum_{k=1}^{n} q_k f (\theta_k, \theta_i) \right) - \sum_{i=1}^{n} q_j p_i (\theta_i, \theta_i) f (\theta_i, \theta_i) > 0,
\]

\[
\sum_{i=1}^{n} q_i f (\theta_i, \theta_i) \left( \sum_{i=1}^{n} q_i p_i (\theta_i, \theta_i) - p_i (\theta_i, \theta_i) \right) > 0,
\]

\(^{10}\) An interpretation is that, if a player wins, they gain only a fixed fraction of the rival’s resources, as assumed, for instance, by Jackson and Morelli (2007)\(^ {18}\).

\(^{11}\) Throughout the paper, we omit, for the sake of space, the dependence of \( f (\theta_i, \theta_j) \) on any variable different than types. Recall, however, that \( f (\theta_i, \theta_j) \) may depend on any number of other parameters, such as, in the examples spelled out so far, the destructiveness parameters \( d_1, d_2 \), and the cost of war \( c \).

\(^{12}\) Note that we allow for \( p (\theta_i, \theta_i) \neq 1/2 \). As pointed out by Jackson and Morelli (2007)\(^ {18}\), “This allows, for instance, to have some geographic, population, or technological advantage or disadvantage.”
and since $\forall i \in \Theta$, there is a $j \in \Theta$ such that $i = j$ and we can take $p_i (\theta_i, \theta_i)$ outside the running sum of $j$s and obtain

$$\sum_{i=1}^{n} q_i f (\theta_i, \theta_i) \left[ \sum_{j \neq i}^{n} q_j p_i (\theta_i, \theta_j) + (q_i - 1) p_i (\theta_i, \theta_i) \right] > 0,$$

$$\sum_{i=1}^{n} q_i f (\theta_i, \theta_i) \left[ \sum_{j \neq i}^{n} q_j p_i (\theta_i, \theta_j) - \sum_{k \neq i}^{n} q_k p_i (\theta_i, \theta_i) \right] > 0,$$

since $q_i - 1 = - \sum_{k \neq i} q_k$. Collecting $q$s, we obtain

$$\sum_{i=1}^{n} q_i f (\theta_i, \theta_i) \left[ \sum_{j \neq i}^{n} q_j (p_i (\theta_i, \theta_j) - p_i (\theta_i, \theta_i)) \right] > 0.$$

In the above expression, the sign of $p_i (\theta_i, \theta_i) - p_i (\theta_i, \theta_i)$ depends on $j \leq i$; thus, we cannot yet conclude the left-hand-side positivity. However, we can rewrite the above expression as follows:\(^{13}\)

$$\sum_{i=1}^{n} q_i f (\theta_i, \theta_i) \left[ \sum_{j=i+1}^{n} q_j (p_i (\theta_i, \theta_j) - p_i (\theta_i, \theta_i)) + \sum_{j=1}^{i-1} q_j (p_i (\theta_i, \theta_j) - p_i (\theta_i, \theta_i)) \right] > 0,$$

$$\sum_{i=1}^{n} q_i f (\theta_i, \theta_i) \left[ \sum_{j=i+1}^{n} q_j (p_i (\theta_i, \theta_j) - p_i (\theta_i, \theta_i)) - \sum_{j=1}^{i-1} q_j (p_i (\theta_i, \theta_j) - p_i (\theta_i, \theta_i)) \right] > 0,$$

$$\sum_{i=1}^{n} \sum_{j=i+1}^{n} q_i q_j f (\theta_i, \theta_j) (p_i (\theta_i, \theta_j) - p_i (\theta_i, \theta_i))$$

$$- \sum_{i=1}^{n} \sum_{j=1}^{i-1} q_i q_j f (\theta_i, \theta_j) (p_i (\theta_i, \theta_j) - p_i (\theta_i, \theta_i)) > 0.$$

For all pairs $(i, j)$ of the first double summation, there exists a unique pair of the second double summation with $i$ and $j$ swapped, and vice versa: e.g., when $(i, j) = (2, 4)$ in the first double summation, there exists an element with $(i, j) = (4, 2)$ in the second double summation, and vice versa. In fact, the number of elements is $n (n - 1) / 2$ in both double summations. Therefore, we can merge the two double summations into the following expression:

$$\sum_{i=1}^{n} \sum_{j=i+1}^{n} q_i q_j \left[ f (\theta_i, \theta_i) - f (\theta_i, \theta_j) \right] \left[ p_i (\theta_i, \theta_j) - p_i (\theta_i, \theta_i) \right] > 0. \quad (4)$$

Since $j > i$, $\theta_j > \theta_i$, and thus by $(MonSpoils)$ and $(MonProb)$, all elements of the summation of Equation (4) are strictly positive, being the product of two strictly negative elements. This proves Equation (4) and, thus, concludes the proof for the second player. The analogous proof works considering the payoffs of the second player of generic $\theta_i$ by simply replacing $q$s with $r$s in the above steps. In fact, the informational naivety or informational sophistication of a player does not affect the other player’s expected payoff and, hence, the other player’s incentive to declare war. \(\Box\)

5. Extensions

More than two players. When players have more than one rival, both $(MonProb)$ and $(MonSpoils)$ can be promptly generalized by redefining $\theta_j$ as the vector of rivals’ types; that is, $i$s probability of

\(^{13}\) Notice that, when $i = n (i = 1)$, the first (second) summation within the square bracket is null.
victory decreases and the spoils of war increase in any rival’s type. Under such generalized assumptions, Proposition 1 carries over. Such a generalization would add extra weight to the notation and would come at a cost of space without adding much insight, and thus, we omit its formal analysis.

**Interior informational naivety.** IN players fully neglect the correlation between rival’s actions and types. A handy concept to capture partial neglect is that of partial cursedness (see Eyster and Rabin, 2005) [11]. Denoting by $\chi \in [0, 1]$, the cursedness parameter, $\chi = 0$ corresponds to IS and $\chi = 1$ corresponds to IN. The expected payoff from war of a player affected by general cursedness $\chi \in [0, 1]$ and of type $\theta_i$ is as follows:

$$
\chi \left[ \sum_{t=1}^{n} q_t p_t (\theta_i, \theta_j) \right] \left[ \sum_{k=1}^{n} q_k f (\theta_i, \theta_k) \right] + (1 - \chi) \sum_{t=1}^{n} q_t p_t (\theta_i, \theta_i) f (\theta_i, \theta_i),
$$

(5)

The expected payoff increases in $\chi$ if and only if Equation (3) holds: hence, players’ expected payoffs from war increase in the cursedness parameter $\chi$, thus generalizing the result of Proposition 1.

**Weakening (MonProb) and (MonSpoils).** We required $f (\theta_i, \theta_j)$ to strictly increase in $\theta_i$ and $p_t (\theta_i, \theta_j)$ to strictly increase in $\theta_j$. This guarantees that every element of Equation (4) is strictly positive. Nevertheless, it suffices that $f (\theta_i, \theta_j)$ and $p_t (\theta_i, \theta_j)$ weakly increase and that both strictly increase for at least one pair of types, so as to have that Equation (4) holds with strict inequality.\(^{14}\)

**Type-conditional expected payoff from war.** In the stylized example of Section 2, the increase due to IN in a player’s expected payoff from war is identical for 0- and R-resource players; that is, $\pi_{IN}^0 - \pi_{IS}^0 = \pi_{IN}^R - \pi_{IS}^R$. However, in the general setting of Section 3, this is not necessarily true. Following the steps of the proof of Proposition 1 leading to Equation (4), we can conclude that $E_{IN} \left[ \pi_{IN}^R \right] - E_{IS} \left[ \pi_{IS}^R \right]$ strictly decreases (increases) in $\theta_i$ if and only if, for all types $t$,

$$
\sum_{i=1}^{n} \sum_{j=i+1}^{n} q_i q_j \left[ f (\theta_{i+1}, \theta_i) - f (\theta_{i+1}, \theta_j) \right] \left[ p_{i+1} (\theta_{i+1}, \theta_i) - p_{i+1} (\theta_{i+1}, \theta_j) \right] > (<) 0,
$$

(6)

However, (MonProb) and (MonSpoils) do not suffice to shed light on the sign of Equation (6). One simple way to do so is by assuming strict supermodularity or submodularity of spoils $f$ and probabilities $p$.\(^{15}\) that is, the marginal benefit (in terms of spoils or probability) of defeating a rival of higher type increases in one’s own type. The stylized example of Section 2, where $f (\theta_i, \theta_j) = \theta_i + \theta_j - c$, is the knife-edge case of modular spoils.\(^{16}\)

### 6. Game-Free Model

The comparison of players’ expected payoffs from war under informational sophistication and informational naivety sufficed to microfound mutual optimism. However, a common approach is to model players’ choice between war and peace, among others, as endogenous and to characterize the equilibria of such a game. In this section, we discuss the pros and cons of the two approaches: our game-free approach and the alternative game-idiiosyncratic approach, where players have choice variables and where equilibria are derived.

---

14 The all-pay auction technology, where $p_t (\theta_i, \theta_j) = 1$ when $\theta_i > \theta_j$, $p_t (\theta_i, \theta_j) = \chi$ when $\theta_i < \theta_j$, and $p_t (\theta_i, \theta_i) = \chi$ when $\theta_i = \theta_i$, satisfies the weakening of (MonProb), but for some pairs of types, it may not satisfy the original (MonProb).

15 The function $f : \mathbb{R}^2 \to \mathbb{R}$ satisfies strict supermodularity in $\langle \theta_i, \theta_j \rangle$ if $f (\theta_i, \theta_j') - f (\theta_i', \theta_j) > f (\theta_i, \theta_j') - f (\theta_i', \theta_j')$ for any $\theta_i' > \theta_i$ and $\theta_j' > \theta_j$. Strict submodularity is similarly defined.

16 Knife-edge as $f (\theta_i', \theta_j') - f (\theta_i', \theta_j) = f (\theta_i, \theta_j') - f (\theta_i, \theta_j)$.
Our game-free approach has a number of advantages. First, the results derived in a game-free model are not an artifact of a particular game form. Second, our game-free model encompassed the concept of informational naivety in players’ expected payoffs, sidestepping the need to characterize equilibria. Third, and most importantly, in our game-free model, we do not need to take a stand on many rather controversial modeling issues. An example is whether the conflict is unilateral or bilateral, that is, whether one side could be forced to fight even if it wished to avoid the conflict or whether the conflict occurs only if both sides choose to stand firm. Examples of models with unilateral conflicts are Powell (1993) [22], Jackson and Morelli (2007) [18], and Slantchev and Tarar (2011) [1] and that with bilateral conflicts are Bueno de Mesquita and Lalman (1992) [23], Fearon (1994) [24], and Fey and Ramsay (2007) [25]. Other examples are the timing of the game, the specific way war is destructive or costly, players’ choice variables (e.g., military expenditure or choice to declare war), or the conflict technology (e.g., whether the probability of victory is proportional to ratios or differences of military efforts).

Nevertheless, working with a game-free model comes at a cost: we cannot derive testable predictions concerning players’ behavior. Superimposing a game to our model would yield more specific and case-by-case predictions and comparative statics. In this sense, a game-idiosyncratic approach would complement our game-free approach.

In the trade-off between the generality of the model and the resulting testable predictions concerning players’ behavior, we opted for the first so as to test the reach of our microfoundation exercise. Nevertheless, in the next section, we analyse a simple game-idiosyncratic model, where players’ efforts are endogenous, and we derive comparative statics on efforts, so as to complement our game-free model.

7. Game-Idiosyncratic Model

While, so far, we focused on informational considerations, in this section, we address strategic considerations more explicitly. In particular, we explicitly model efforts denoted by \( e^i_j \) for player \( i \in \{1, 2\} \) with resources \( j \in \{0, R\} \) as a choice variable, and the individual probability of victory is modeled à la Tullock; that is, player \( i \)'s probability of victory equals their own effort divided by the total effort.\(^{17}\) Both under IS and IN, we analyse type-symmetric equilibria; that is, if the two players have the same level of resources, they exert the same equilibrium level of effort—i.e., \( e^0_1 = e^0_2 = e^0 \) and \( e^R_1 = e^R_2 \equiv e^R \).

In this setting with endogenous effort and Tullock conflict, we maintain the fixed costs paid by the winner as in the model of Section 2 (i.e., \( f(\theta_i, \theta_j) = \theta_i + \theta_j - c \)) but we additionally assume that, regardless of the outcome of the war, each player pays a cost of effort equal to the effort exerted by that player.

7.1. Informational Sophistication

Under IS, the expected payoffs from war of a player with resources \( R \) and 0, respectively, equal\(^{18}\)

\[
\frac{1}{2} \left[ \frac{e^R_1}{e^R_1 + e^R_2} (2R - c) + \frac{1}{2} \frac{e^R_1}{e^R_1 + e^R_2} (R - c) - e^R_1 \right] \quad \text{and} \quad \frac{1}{2} \left[ \frac{e^0_1}{e^0_1 + e^0_2} (R - c) + \frac{1}{2} \frac{e^0_1}{e^0_1 + e^0_2} (-c) - e^0_1 \right].
\]

If \( c = 0 \), routine maximization steps yield the unique type-symmetric equilibrium:

\[
e^0 = \frac{-1 + \sqrt{5}}{16} \quad \text{and} \quad e^R = \frac{3 + \sqrt{5}}{16}. \tag{7}
\]

\(^{17}\) If both players exert 0 effort, each player has 1/2 probability of victory.

\(^{18}\) Throughout this subsection and the next one, we spell out the maximization problem of player 1. That of player 2 is symmetric.
If \( c \in [R, 2R] \), we immediately obtain \( e^0 = e^R = 0 \) from the maximization problem of a player with resources 0, and plugging this result into the FOC of a player with resources \( R \) gives the following unique type-symmetric equilibrium:

\[
e^0 = 0 \text{ and } e^R = \frac{2R - c}{8}.
\] (8)

For intermediate values of \( c \), obtaining a closed-form solution for equilibrium efforts is challenging, and in fact, there are two interior type-symmetric equilibria. However, even without closed-form solution, one can easily prove that, in any interior equilibrium, the following property holds.

**Lemma 1.** If \( c \in (0, R) \), in an interior type-symmetric equilibrium of the game under IS, \( e^R > e^0 \).

**Proof.** Assume by contradiction that \( e^0 \geq e^R \). The FOCs for the two types of player 1 read:

\[
\begin{align*}
\text{FOC}^R & : \frac{1}{2} \left( \frac{e^R}{e^R + e^0} \right)^2 (2R - c) + \frac{1}{2} \left( \frac{e^R}{e^R + e^0} \right)^2 (R - c) = 1, \\
\text{FOC}^0 & : \frac{1}{2} \left( \frac{e^0}{e^0 + e^R} \right)^2 (R - c) + \frac{1}{2} \left( \frac{e^0}{e^0 + e^R} \right)^2 (-c) = 1,
\end{align*}
\]

and the FOCs for the two types of player 2 are symmetric. Applying type-symmetry, we obtain

\[
\begin{align*}
\text{FOC}^R & : \frac{1}{8e^R} (2R - c) + \frac{1}{2} \left( \frac{e^0}{e^R + e^0} \right)^2 (R - c) = 1, \\
\text{FOC}^0 & : \frac{1}{2} \left( \frac{e^R}{e^R + e^0} \right)^2 (R - c) + \frac{1}{8e^0} (-c) = 1,
\end{align*}
\]

and considering the ratio of the addends containing the term with \((R - c)\) in both equations, we obtain

\[
e^0 \frac{e^R}{e^R} = \frac{1}{1 - \frac{1}{8e^0} (-c)}.
\]

By \( e^0 \geq e^R \), the numerator of the right-hand side (RHS) has to be greater than the denominator of the RHS, or equivalently

\[
\frac{2R - c}{e^R} < \frac{-c}{e^R},
\]

which is a contradiction. \( \square \)

**7.2. Informational Naivety**

Under IN, the expected payoffs from war of a player with resources \( R \) and 0 respectively equal

\[
\left( \frac{1}{2} \frac{e^R}{e^R + e^0} + \frac{1}{2} \frac{e^R}{e^R + e^0} \right) \left( \frac{3}{2} R - c \right) - e^R \text{ and } \left( \frac{1}{2} \frac{e^0}{e^1 + e^R} + \frac{1}{2} \frac{e^0}{e^1 + e^0} \right) \left( \frac{1}{2} R - c \right) - e^0.
\]

If \( c = 0 \), routine maximization steps yield the following unique type-symmetric equilibrium:

\[
e^0 = \frac{7R}{64} \text{ and } e^R = \frac{21R}{64}.
\] (9)

---

19 Throughout this subsection and the next one, the second-order conditions (SOCs) hold.
If $c \in [R/2, 3R/2]$, we immediately obtain $e^0 = e^R = 0$, and plugging this result into the FOC of a player with resources $R$ gives the following unique type-symmetric equilibrium:

$$e^0 = 0 \text{ and } e^R = \frac{3R - 2c}{16}. \tag{10}$$

For intermediate values of $c$, obtaining a closed-form solution for equilibrium efforts is less challenging than under IS and the type-symmetric equilibrium is unique. However, its derivation is routine and space consuming, and thus, we only write here the final result.

$$e^0 = \frac{8c^2 - 16cR + 7R^2}{64 (R - c)^2} (R - 2c) \quad \text{and} \quad e^R = \frac{8c^2 - 16cR + 7R^2}{64 (R - c)^2} (3R - 2c). \tag{11}$$

Hence, the result corresponding to Lemma 1 (i.e., $e^R > e^0$) immediately follows from the comparison of the two above expressions.\(^{20}\)

### 7.3. Comparative Statics

It is straightforward to verify the result of Proposition 1 (proved in a game-free setup) by plugging the equilibrium efforts back into individual payoffs. However, the advantage of a game-idiosyncratic model is the possibility of sharper predictions and comparative statics.

First, for every value of $c \geq 0$ and regardless of IS or IN, the equilibrium effort of a player strictly increases in their own resources (i.e., $e^R > e^0$), and consequently (MonProb) follows. This can be immediately verified in (7), (8), and Lemma 1 for IS and in (9), (10), and (11) for IN.

Second, we can compare for a given level of resources whether a player’s effort is greater under IS or IN. If $c = 0$, a simple comparison of (7) and (9) shows that a player’s effort is smaller under IS than under IN. If $c \in [R, 3R/2]$, a simple comparison of (8) and (10) shows that a player’s effort is greater under IS than under IN. If $c$ takes intermediate values (i.e., $c \in (0, R)$), the lack of tractability and multiplicity of equilibria in the case of IS does not allow an easy comparison.

The intuition why efforts could be greater under IS or under IN is as follows. On the one hand, a player affected by IN overestimates their expected payoff from war relative to IS predictions, and hence, they would a priori exert more effort under IN than under IS. On the other hand, a game-idiosyncratic approach makes a player affected by IN anticipate that their rival, being also affected by IN, expects a high payoff, and this discourages the first player’s effort. The former effect turns out to be stronger when $c \in [R, 3R/2]$ (i.e., a player’s effort is greater under IS than under IN), while the latter turns out to be stronger when $c = 0$ (i.e., a player’s effort is greater under IN than under IS). This tension between opposing forces is clearly possible only when efforts are endogenous, which is the novelty of this section.

### 8. Discussion

Mutual optimism about the outcome of war is typically exogenously imposed in models on rationalist explanations for war. In this paper, we borrow from the economic literature a condition on the information processing capabilities of country leaders and show that such a condition gives rise to mutual optimism. Thus, in contrast to the canonical approach discussing mutual optimism as a source of outbreak of war, we take a step back and highlight a source of mutual optimism itself. We identify informational naivety as a source of mutual optimism: namely, country leaders fail to understand the correlation between other players’ actions and their types despite understanding everything else in a conventional game theoretic sense, including the correct distribution of other players’ actions and their types. We benchmark the setting with informational naivety to the one without (i.e., informational

---

\(^{20}\) Steps identical to the proof of Lemma 1 would also show that $e^R > e^0$ in any interior type-symmetric equilibrium.
sophistication) and find that informational naivety unambiguously increases players’ expected payoffs from war.

Interpreting players’ expected payoffs from war as a proxy for the likelihood of war, as intuition and the literature suggest,\(^\text{21}\) enhances the scientific payoff of the present exercise. Nonetheless, we remained agnostic about the specific channel through which larger expected payoffs from war increase the probability of war: for instance, larger expected payoffs from war could directly trigger an attack or could make it harder to reach an agreement that both players perceive as mutually advantageous.

It is important to compare our informational naivety to the canonical mutual optimism about the likelihood of victory, where subjective probabilities of victory sum to more than one.\(^\text{22}\) To grasp the intuition, Fearon (1995)\(^\text{[5]}\) considers a situation where two states bargain over the division of $100 and each has the outside option of going to war. If each expects that it surely would prevail at war, then each side’s expected value for the war option is $80; as in Fearon’s model, going to war entails a fixed cost of $20. Therefore, given these expectations, neither side will accept less than $80 in the bargaining, implying that no negotiated outcome is mutually preferred to war. More generally, suppose that state A expects to win with probability \(p\), state B expects to win with probability \(r\), and \(p + r\) sum to greater than one. Such conflicting expectations will certainly shrink and could eliminate any ex ante bargaining.

Both Section 2’s and Fearon’s examples depict noncomplex environments where irrational players compare in their minds war–peace payoffs. The two irrationalities are different in nature, and hence, the justifications proposed for the canonical mutual optimism do not automatically justify informational naivety. The canonical mutual optimism is traditionally, and still nowadays, justified with overconfidence: moods which cannot be grounded in fact result in a process by which nations evade reality (Blainey, 1988)\(^\text{[3]}\). Overconfidence has been proved to have strong predictive power in a number of fields other than conflict analysis. Informational naivety is also grounded on a single psychological principle, which, instead of being overconfidence, is the underappreciation of the informational content of other people’s behavior (Eyster and Rabin, 2005)\(^\text{[11]}\).

The canonical argument in support of informational naivety in fields other than conflicts is as follows. When augmenting a theoretical model by IN, its analytical result is typically checked to be consistent with the empirical observation. For instance, in auctions, the analytical result is that bidders overbid, which is in line with empirical findings. This empirical exercise typically sufficed as supportive evidence for at least two reasons: (i) the first-best direct observation of beliefs and, hence, miscalculations in line with IN is hardly available, and hence, empirically verifying the theoretical prediction is second-best, and (ii) IN is based on a well-established psychological principle. Clearly, points (i) and (ii) carry over to any setting with IN other than auctions and, hence, also to conflicts. Thus, similarly to auctions, we notice that the theoretical predictions of the model with IN (excessive bidding due to overestimation of the value of the object auctioned off conditional on victory or excessive war declarations due to the overestimation of the payoff from war) are consistent with real-world observation, missing the first-best direct evidence of IN beliefs. Finally, IN has similarly been proven to have significant predictive power in a number of fields other than conflict analysis: in fact, the predictions of informational naivety has been shown to be consistent with several real-life anomalous

\(^{21}\) For pioneering works, see, for instance, Bueno de Mesquita and Lalman (1986)\(^\text{[26]}\), Morrow (1989)\(^\text{[4]}\), and Banks (1990)\(^\text{[27]}\).

\(^{22}\) A similar argument could be made for a systematic overestimation of the spoils of war: subjective estimations of the spoils of war sum to more than the objective stakes.
phenomena that conventional informational sophistication fails to capture, such as winner’s curse in auctions, various herd behaviours, and trade in markets with adverse selection.

Similarly to IN, the behavioural assumption of overestimation of likelihood of victory also suffers from missing any direct empirical evidence despite its theoretical prediction of excessive war declaration being widely observed. The reason is once again that directly observing and measuring such beliefs of country leaders is challenging. In fact, we are not aware of any empirical evidence that country leaders systematically overestimate the likelihood of victory. Nevertheless, as for our irrationality, scholars provided direct evidence that country leaders may not process information in an informationally sophisticated way: classical evidence of errors in processing information comes from the psychological international relations literature—in particular, see Jervis (1976) and Jervis, Lebow, and Stein (1985). As pointed out by Fey and Ramsey (2007), country leaders who have many responsibilities may face a volume of information that induces flaws in their learning.

We conclude with two remarks. First, informational naivety has been typically applied to rather complex strategic and informational environments, while we apply it to a simple correlation between two exogenous variables. However, recall that superimposing a game to our model, for example, endogenizing efforts as we did in Section 7, would not change the main result of overestimation of war payoffs as long as \((\text{MonProb})\) and \((\text{MonSpoils})\) hold. For this reason, existing evidence of informational naivety from complex environments encompasses our noncomplex environment too. Second, despite the high stakes of international conflicts instinctively suggesting that country leaders meticulously assess costs and benefits of a conflict, informational naivety has been successfully put forth as an explanation for failures in other high-stakes environments. For instance, when firms decide whether to enter a market, a firm neglecting that only firms sufficiently skilled decide to enter would overestimate their own profits from entry and, hence, over-enter: such a reference-group neglect generated by informational naivety could hence explain why most new businesses fail shortly after entry (Camerer and Lovallo, 1999). Inflated credit ratings are another high-stakes environment where, as pointed out by Skreta and Veldkamp (2009), informational naivety could explain the upward bias in the rating of structured credit products, which is widely cited as one contributor to the crisis.

**Funding:** This research received no external funding.

**Acknowledgments:** I thank the Editor, Galina Zudenkova, and two anonymous referees for helpful comments and suggestions. I also thank Bharat Goel, Raphaela Hennigs, Kai Konrad, Jonas Send, and the participants at the XXXII Tax Day for useful discussion. All errors are my own.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


---

23 When bidders share a common but unknown value for the object at auction and receive private signals about such value, they tend to bid more than equilibrium theory predicts; this phenomenon is known as the winner’s curse. A bidder wins if others bid sufficiently lower than them, which happens if others received private signals that are more negative than their own. A bidder who fails to draw such an inference between others’ bids (actions) and signals (types) overestimates the value of the object and, hence, overbids (e.g., Kagel and Levin, 1986; Eyster and Rabin, 2005; and the references therein).

24 Subjects disproportionally enter competitions on easy tasks (Moore and Cain, 2007). eBay sellers disproportionately choose to end their auctions during the times of day when more bidders are online (Simonsohn, 2010). Informationally naive people herd with positive probability on incorrect actions (Eyster and Rabin, 2010).

25 If a prospective buyer fails to realize that a bid will only be accepted if the seller’s valuation is less than the bid, then the buyer might incur a loss (e.g., Holt and Sherman, 1994).

26 See Lindsey (2018) for a recent contribution overcoming the issue of the difficulty of observing country leaders’ beliefs. See also Lai (2004) and Bas and Schub (2016) for similar empirical evidence. For laboratory experiments directly manipulating information, see Tingley and Wang (2010) and Querk (2015).
20. Fudenberg, D. Advancing Beyond Advances in Behavioral Economics. J. Econ. Lit. 2006, 44, 694–711. [CrossRef]
21. Jackson, M.O.; Morelli, M. Political Bias and War. Am. Econ. Rev. 2007, 97, 1353–1373. [CrossRef]
40. Lindsey, D. Mutual Optimism and Costly Conflict: The Case of Naval Battles in the Age of Sail. J. Polit. 2018, 81, forthcoming. [CrossRef]
