Uncertainty, Overconfidence, and War

Maxime Menuet, Petros Sekeris

Uncertainty, Overconfidence, and War☆

Maxime Menuet*, Petros G. Sekeris**

Abstract

The present paper studies the causes and duration of wars by building a war of attrition game, and explores the effect of overconfidence in such settings. During the fight, each player infers his opponent's inclination in surrendering given two psychological biases jointly capturing overconfidence: illusory superiority (overestimation), and over-self-confidence (overprecision). We demonstrate that overconfidence is neither necessary, nor sufficient to have war. Yet, overconfident decision-makers are nevertheless more likely to initiate war, and to remain active longer in a conflict. Moreover, we show that the effect of overestimation on war duration may be non-monotonic, with the duration of wars increasing in overconfidence for lowly overconfident players, and decreasing for highly overconfident ones. We argue that this simple model helps understanding a host of real-world conflictive situations.

JEL: D74, C72, D82

Keywords: Overconfidence, Imperfect information, War of attrition, Illusory superiority

1. Introduction

Informational problems have often been identified by international relations scholars as a central factor explaining wars. From events dating back to the Peloponnesian

*CERDI, UMR 6583, CNRS, IRD, Université Clermont-Auvergne e-mail: maxime.menuet@uca.fr

**Montpellier Business School, 2800 Avenue des Moulins, 34080, Montpellier, France. e-mail: p.sekeris@montpellier-bs.com

War (431-404 BC) that opposed Athens to Sparta to more recent ones as the Operation Iraqi freedom, contenders have repeatedly mis-estimated their capacity to militarily achieve desired outcomes (e.g. Levy 1983, Lake 2010), eventually giving rise to the outbreak of wars, but also in turn explaining their duration. Upon a closer look, however, one can distinguish two types of informational problems: rational, and non-rational ones. Yet these two types of informational problems have not received the same attention by the scholarship: a plethora of theoretical models have attempted understanding the impact of rational informational biases on conflict (see Bas and Schub for a recent review), while few theoretical advances have been made to date to understand the implications of irrational biases in international relations (e.g. Wittman 1979, Stam and Smith 2004), or how these two types of informational problems interact to explain the roots and duration of war. Our focus in this article is on this latter type of informational problem.

There exist several different non-rational informational biases, which all share the attribute of distorting one's believed expected probability of an event away from the (true) expected probability of that realization. In the specific context of war, these include biased beliefs on one's adversary's intentions or capabilities, or of his capacity to accurately evaluate one's own intentions or capabilities (e.g. Levy 1983). In this article we restrict our attention to excessive optimism, or overconfidence, which in conflictual situations has frequently been identified as a major cause of wars (e.g. Levy 1983, Blainey 1988, Jervis 1988, Johnson 2004).[1] The argument is rather straightforward: if a country (mistakenly) believes it is militarily stronger than its opponent, then when the latter correctly estimates its own winning odds, or if it is similarly biased, the peaceful bargaining range evaporates with war potentially resulting. Overconfidence is also typically associated with lengthy wars of attrition where overconfident contenders refuse to surrender even when the odds are at their disadvantage, pursuing combat in the hope of seeing their adversary backing down. The duration of World War I, of the Vietnam war, or of the Russian invasion of Afghanistan, for instance, can partially be explained by overconfident parties eventually losing the war (Germany, the U.S.A., and Russia, respectively). These lines of reasoning may appear sound, but they are nevertheless incomplete as we demonstrate in this article.

We propose an original attrition warfare which is well suited to explain both the onset and duration of wars. Attrition warfare is a useful theoretical framework to model long-lasting conflicts since it conceptualizes war as a costly endeavor to all participants, with the victor being the last man standing still after his foes have dropped out of the contest. Several reasons make a model of attrition suitable for studying the duration of wars (Powell 2017) and, by extension, the decisions to go

[1] Rational choice theorists have identified a vast list of causes of war, including information asymmetries (Bueno de Mesquita 1981, Fearon 1995), commitment problems (Powell 2006), or political biases of decision makers (Jackson and Morelli 2007).

to war. Levy (1983) highlighted that wars that are initiated are wars that at least one contestant expects to win, but the same reasoning applies to the decision to end a war. If the opponent expected a better outcome by putting an early end to the military confrontation he would have acted accordingly. Hence, one is led to deduce that the non-termination of wars, i.e. their continuation, is an endogenous decision which is well captured by a war of attrition model. Such a setting is therefore particularly appealing for modeling interstate conflicts since these quite often end with the capitulation of one side.

In our attrition model two countries vie for the control of each other's wealth and each side decides (i) whether to declare war, and (ii) when to step out of an ongoing war, i.e. when - if ever - to declare defeat. Leaders have imperfect information on the resilience of their opponent's wealth broadly defined, namely on the damage their opponent's resources sustain in war, and on the degree the local populations' post-war support for the foreign conqueror is dampened by on-going combat.[2] Moreover,

[2]Shirkey (2016) develops several arguments as to why imperfect information on the contestants' resilience explains wars' initiation and duration, but also why such information asymmetry is not resorbed on the battlefield. On the one hand, one could argue that such resilience solely affects the fighting incentives of the country initially in control of a territory in case it is victorious (i.e. more costly wars push actors to yield easier). On the other hand, the value of the conquered bounty is equally decreasing in the war's length for the foreign conqueror both because of the physical damage inflicted on the valuable resources, and because the "hearts and minds" of local populations become increasingly alienated (Byman 2006, Lyall et al. 2013), thereby resulting in higher occupation costs. This well exemplified by the operation *desert storm* in 1991. After Saddam's forces invaded Kuwait, the U.S. senate approved a military intervention, and yet after after "the US military expelled Iraq from Kuwait, [it] ended the fighting prematurely" (Ricks, 2006). Several geopolitical reasons have driven such decision, including maintaining a balance of power in the Middle East. Yet, as Saddam's troops were retreating from Kuwait, they set on fire the oil fields, and it is not far fetched to conclude that the US rested on a partial victory instead of engaging a costly war that would have annihilated Iraqi's economy and thereby yielded few benefits to the US. In Bush's own

the two leaders may be overconfident in their estimation of the damage inflicted on their opponent's wealth (overestimation) and may thus overestimate the cost of war to their opponent, as well as the damage operated on local populations' "hearts and minds". Second, they may be overly optimistic in the confidence they have in their prediction of their opponent's resilience (overprecision). Our model predicts that in the absence of overconfidence, peace and war equilibria co-exist. While peace is expected to emerge given the costliness of fighting, war is also an equilibrium because of a reasoning specular to the well-known risk-reward trade-off. Overconfidence biases decision makers towards initiating war, and yet the peace equilibrium always subsists by the same risk-reward type of reasoning. Regarding the war's durability, our findings are richer. Based on their imperfect information of their opponent's resilience, and on their subjective evaluation of that information, contenders evaluate their optimal time of dropping out of the conflict by comparing at any time period the marginal (expected) return from pursuing fighting, with the marginal cost. As with war initiation, and for the same reasons, the drop-out time can be strictly positive or immediate. When focusing on equilibria with a strictly positive drop-out time, however, while overconfidence could be expected to lengthen inefficient wars, this intuition ought to be nuanced. When increasing the overestimation bias of a leader on

words, "I firmly believed that we should not march into Baghdad.... It would have taken us beyond the imprimatur of international law... assigning young soldiers to a fruitless hunt for a securely entrenched dictator and condemning them to fight in what would be an unwinnable urban guerrilla war." Bush and Scowcroft (1998).

the expected damage incurred by its opponent, that decision-maker accordingly updates downwardly the expected drop-out time of his foe, thus making oneself willing to remain active in the conflict for a longer time-period. Yet, this same expectation reduces the expected value of victory since with the local populations' "hearts and minds" being more alienated, the worth of the contested territory is downwardly biased. The optimal drop-out time is a combination of these two mechanisms, and we show that the second mechanism can outweigh the first one for sufficiently high levels of overconfidence.

The workhorse model for modeling war has been the bargaining model that gained a lot of popularity with the contributions of Fearon (1995) and Powell (2006). While bargaining is well suited for understanding the roots of wars, it is much less informative on their duration, unless coupled with dynamically changing features such as relative capabilities or resolve (e.g. Powell 2012). In our model the information held by contenders on one another's resilience may be perfectly informative on the objective features of resilience, which would suggest in a dynamic bargaining settings the absence of wars. And yet, because of (unresolvable) uncertainty, and because of overconfidence, we demonstrate that wars may nevertheless take place, and we are equally able to characterize their duration. Attrition models have seldom been used in the international relations literature, with Fearon (1998) and Powell (2017) being two notable exceptions. Fearon (1998) proposed a model of (international) bargaining in a dynamic setting where costly delay takes place in equilibrium. Of closer interest to our own study is the article of Powell (2017) who develops a war of attrition model involving two contenders and a third party allowed to support either side. As contenders only know the distribution of their opponent's type, and do not know whom they are facing with certainty, the support of the third party for either side will incentivize "weak" types of the supporting country to remain active longer in the war of attrition, eventually lengthening the expected duration of war. Since, however, the third party is aiming for a short war, at equilibrium it will "mix" its support to the two contenders to disrupt the above mechanism. Our article is therefore the first contribution to explain war initiation and its duration with an attrition model, which also implies that beyond our theoretical findings proper, we equally propose a new toolset for international relations scholars.

The next section presents an overview of the literature and highlights the importance of introducing overconfidence in the analysis. Section 3 develops the theoretical model, that we solve in Section 4, before exploring the effect of overconfidence in Section 5. In Section 6 we briefly discuss applications of our theory in the context of the WWI, the Peloponnesian Wars, and the Vietnam War. Lastly, Section 7 concludes.

2. Literature Review

In this section, we first present the literature on imperfect information and war, before reviewing existing advances on the relationship between overconfidence and war.

Imperfect information and war

The role of imperfect information in explaining war onset and war duration has received extensive attention in the literature, and we sequentially overview the existing theories and debates.

Early on, Blainey (1988) argued that war can ensue out of mutual optimism: "war is usually the outcome of a diplomatic crisis which cannot be solved because both sides have conflicting estimates of their bargaining power" (Blainey 1988: 114). Fearon (1995) further developed this rationalist argument for why imperfect information can be conducive to war by viewing war as the failure to reach a negotiated bargained agreement. War is costly, and bargained solutions must exist that leave all contenders better-off than by fighting. Yet, if contenders are endowed with private information on elements determining their likelihood of prevailing on the battlefield (e.g. weaponry, troops' morale, political support etc. . .), and given the cost of disclosing private information and the incentives of "weak" players to appear "strong", this may fall short of becoming public. With the information remaining private, the *casus belli* demands of the contenders may be biased upwards and will then sum up to more than the aggregate value contenders would be sharing under a bargained agreement. Contenders will then probabilistically engage in conflict in what has been termed the *risk-reward trade-off*: one risks having a higher demand rejected and war taking place, rather than increasing the probability of a negotiated settlement with a less favorable peaceful agreement.

A vast literature on the topic has emerged over the past 20 years, with much emphasis given on bargaining models, but not exclusively (e.g.- Powell 1996, or Ramsay (2017) for a literature review). With everything fixed but the decision to go to war or not, Fey and Ramsay (2007) argue that mutual optimism in the absence of communication between contenders cannot be conducive to war. Yet, their argument relies on the assumption that war can only result if both opponents opt for it, a definitional nuance picked up by Slantchev and Tarar (2011) who subsequently showed that the risk-reward trade-off is a reality if two rational enemies that do not bargain can provoke war unilaterally; a result echoing Baliga and Sjöstöm (2008) who demonstrate why weak contenders can have incentives to pretend to be strong to avoid the confrontation in the first place.[3] The element that glues together this literature and echoes the earlier literature is undeniably that imperfect information can explain war, with one decision-making protocol or another.[4]

Irrespective of the modeling approach used, a consensus exists in the literature: if war is destructive and information is public, there should be no wars.[5] Powell (2006)

[3]The debate seems to have ranged on, however, since a core finding of Slantchev and Tarar (2011) that mutual optimism is both a necessary and sufficient condition for having war, was later questioned by Fey and Ramsay (2016).

[4]Adam and Sekeris (2017) constitutes an exception since they demonstrate that imperfect information can help sustaining the peaceful status quo when contenders are not seen as unitary actors, but where instead each side is subdivided in two players who communicate imperfectly and who jointly decide their country's strategy.

[5]With an endogenous arming process, De Luca and Sekeris (2013) show that there can be no peaceful pure strategy equilibrium since contenders always find it (probabilistically) optimal to increasing their armament level and attack their potentially lowly militarized foe. Similarly, Debs and Monteiro (2014) derive such mixed strategy equilibria in the context of shifting relative power between contenders.

argues that assuming information asymmetries subsist in environments where contenders exchange information in various ways (diplomacy, arms build-up & parades, intelligence reports, military skirmishes, etc. . .) may appear as a strong assumption. Or as Kydd (2005) stated "if uncertainty is at the heart of crises, then communication is the key of resolving them (Kydd 2005: 186). But such information revelation may be costly, as war may be a necessary step to establishing peace. Wagner (2000), Filson and Werner (2002), Slantchev (2003), and Powell (2004), for instance, consider (real) wars and sequential bargaining as confrontations through which contenders reveal one another their relative power through a costly process involving fighting and peace proposals, eventually allowing them to settle on a bargained solution. Thus, while conflict may be sparked by uncertainty, parties will eventually settle as uncertainty gradually disappears. This is not to say information is perfect at any point in time, but rather that contenders should update their information set through time using Bayes' rule, eventually converging to the same information set, or as Aumann (1976) famously stated, eventually "agreeing [on their posterior beliefs] to disagree [on their private information]".

Information revelation needs not occur on the battlefield (Leventoglu and Tarar 2008), but when it does information imperfections can also explain the duration of wars (e.g. Filson and Werner (2002), Slantchev (2003)), even in such instances where contenders do not have common prior beliefs, i.e. when they know that the expected sum of winning probabilities may be different to unity (Smith and Stam

2004). As a consequence the duration of wars is a positive function of the amount of informational asymmetries and negative one of the speed of information revelation. The flip side of this observation, however, is well summarized by the following quote: "The information revelation mechanism cannot deal very well with long wars: to think that it takes many years of near constant interaction for opponents to learn enough about each other is surely stretching the theory" (Leventoglu and Tarar, 2007:756).

In summary then, the literature on imperfect information and war has greatly advanced our understanding of the topic, but appears nevertheless unable to explain why conflicts occur when the contenders' intel' is reliable, or why conflicts often last disproportionally long.[6]

Overconfidence and war

Informational biases may also be irrational, what is commonly known in political science as misperceptions (Jervis 1976). Despite the dearth of theoretical contributions on misperceptions and war (e.g. Wittman 1979, Stam and Smith 2004), historians, political scientists, and psychologists alike remind us of the importance of these psychological biases in explaining war (see, e.g., Levy 1983, Blainey 1988, Van Evera 1999, or LeShan 2002).[7] One specific such bias is overconfidence which has been recognized as being particularly salient in explaining why countries go to war (John-

[6]For alternative theoretical explanations of war durations see e.g. Leventoglu and Slantchev (2007) or Powell (2012).

[7]More generally speaking, a consensus has emerged that perceptual biases can explain what appears as irrational human behaviour (e.g.- Kahneman and Tversky 1996).

son 2004). Overconfidence is a character trait which is more preponderant in leaders than in the rest of the population, and can take various forms in the context of war, including overestimating one's own capacity to defeat an enemy, overestimating one's capacity to hurt the enemy, underestimating the adversary's capacity to sustain war, i.e. his resilience, or any other such bias making a contender believe in his superiority in a war context, and therefore in is higher-than-real odds of winning. A corollary of such individual overestimation of winning probabilities is that if at least some player is overconfident, then the aggregate expected winning probabilities of players will sum up to more than unity, thereby giving rise to war (or its continuation) as in Blainey's *mutual optimism* concept.

Interestingly, however, and in stark contrast with "rational" information biases, overconfidence is not bound to disappear in the presence of perfect information. An overconfident individual will by definition remain fairly insensitive to information on the state of the world, since it would subsist even if the true (expected) was common knowledge. Consider, for instance, an overconfident gambler who believes he has better than fair odds of getting the desired number when rolling an unbiased dice. While the probability of a specific number is commonly known (including to the gambler) to be 1/6, the overconfident player's biased beliefs remain unaffected by recurrent throws of the dice.[8]

[8]For individuals to exhibit overconfidence, some uncertainty, yet not necessarily asymmetric information, on the final outcome needs to be present. If an event is certain, as with such a fact that all humans die at some point, one cannot be overconfident to be exempt from this rule. What

Many psychologists and neuroscientists have long shown that overconfidence is a natural human trait.[9] This psychological tendency describes a cognitive bias where someone subjectively believes that his or her judgment is better or more reliable than it objectively is. The psychological literature (following Pulford, 1996; Moore and Healy, 2008; Ortoleva and Snowberg, 2015) distinguishes two types of overconfidence. Most leaders show a cognitive bias towards: (i) exaggerated individual talents of fighting, called overestimation or illusory superiority that reflects the tendency to believe that we fight better than others; (ii) exaggerated self-confidence, called over-precision that reflects an excessive degree of "optimism" in victory.[10] This judgment bias appears when leaders over-believe in the precision of their estimated victory, and is often linked to a narcissistic psychological trait. According to Stoessinger (2011), for instance, certain wars are due to the political leaders' arrogance, stupidity, carelessness, or weakness, and to the disproportionate weight given to their egos, which explain their overconfidence in victory. The full explanation of the psychological mechanisms underlying overconfidence, and why it can be reinforced on the eve of wars is elaborate and well summarized in Johnson (2004) and Johnson and Tierney (2006). What we can take at face value, however, is the fact that many leaders tend

one can nevertheless be overconfident of is the probability of reaching an old age.

[9]See, e.g., Taylor and Brown (1988, 1994); Peterson (2000, 2006); Erhlinger et al. (2016).

[10]The literature exhibits a third type of overconfidence (the overestimation). In our paper, the over- placement and the overestimation describe the same phenomenon: the player under-estimates the resilience of the opponent, or over-estimates his own resilience to conflict. This modeling is consistent with the literature suggesting that overestimation and overestimation are interchangeable manifestations of self-enhancement (Kwan et al., 2004).

to be overconfident, and it is thus crucial to understand the theoretical mechanisms tying overestimation and overprecision to war and its duration.

In the same vein that mutual optimism can be a catalyst for wars, so can overconfidence. The underlying mechanism driving conflict is the same in that it is a (mistaken) evaluation of one's own prospects of victory - or "unduly rosy estimates of relative military power" (Van Evera 1999: 16) that leads to war in both cases, with the important difference that, unlike overconfidence, mutual optimism is bound to disappear as new information is acquired by the contenders.

Overconfidence also helps us understand wars' duration (Wittman 1979). Indeed, a remarkable feature is that the conflicts with highly overconfident leaders are mainly long-lasting wars (Jervis, 1976; LeShan, 2002). History is replete with examples of long wars that were initially predicted to be short, including the Peloponnesian War (Hanson 2005), the two World Wars (Johnson 2004), or the invasion of Iraq (Sullivan 2007). The ex-post information we have access to reveals a clear overconfidence of the various players involved in these conflicts, especially in light of the intelligence these decision makers deliberately decided to ignore at that time (Johnson 2004, Hanson 2004).

3. A War of Attrition Model

3.1. The fundamentals

We consider a war of attrition with two players indexed by $i \in \{1, 2\}$, each controlling some resource $R_i(t)$ at time $t \geq 0$. The initial value is, for simplicity, normalized to be one, i.e. $R_i(0) = 1$[11]. This resource reflects the value of controlling the said territories, and it therefore captures a combination of the territories' physical resources, alongside the popular support of the local population, its morale, and other relevant features.

The game proceeds in a war pre-stage game, that is followed by a war of attrition if either player declares war. In the pre-stage each player $i = \{1, 2\}$ takes an action $s_i \in \{p, w\}$ and if either player chooses w a war of attrition ensues, otherwise the status quo is preserved. If both players opt for p, the players receive a symmetric payoff $V_i^S = 1$. If, however, either or both players choose $s_i = w$, a war of attrition is initiated and at each instant $t \geq 0$ both players simultaneously decide whether to exit or to continue the war. Remaining active is costly to both players, and we assume that this war of attrition involves two costs. *First*, players incur a cost of the conflict at each period of time irrespective of the war's outcome. We assume without loss of generality that this cost is unitary. *Second*, war directly degrades the value of both players' contested resource because of the physical destruction incurred, alongside the loss of morale of the population, and the subsequent damage to the players' popularity.[12] Accordingly, player i sustains an exogenous cost of war $X_i(t)$ at time t,

[11] Considering asymmetric status quo distributions of resources would not modify the essence of the model's mechanisms, but it would significantly complexify the analysis.

[12] Prolonged wars involve several costs, including physical and human capital destruction, trade disruptions, reduced FDI, and declining tourist activity (e.g. Arunatilake et al., 2001; Abadie and Gardeazabal, 2003; Ben Bessat et al. 2012). On top of these costs, populations typically grow weary with the length of wars, thereby implying that even if eventually victorious, a warring party may nevertheless be losing the battle for the population's hearts and minds as the duration of a war

and his rent can be seen as being degraded as follows:

$$\dot{R}_i(t) = -X_i(t),$$

where $\dot{R}_i(t)$ denotes the time derivative.

For stake of clarity, we assume that $X_i(t) = v_i R_i(t)$, where v_i is the depreciation rate of player i's territory due to the conflict. This parameter encapsulates all factors negatively affecting the rent's value, and we assume it to be constant over fighting time. Accordingly, we can deduce that $R_i(t) = \exp(-v_i t)$.

The value v_i is independently drawn from the well-defined distribution $F_i(\hat{v}, \gamma^2)$ defined over $[\underline{v}, \overline{v}] \subseteq [0, 1]$, where both player share the same mean $(\hat{v})$ and variance (γ^2). We denote by f_i the associated density function, with $f_i(\underline{v}) > 0$ and $f_i(\overline{v}) > 0$. In an imperfect information setup, player i only knows the realization of his type, and has beliefs about the opponent's type whose distribution, $\tilde{F}_{-i}$, depends on overconfidence parameters as detailed below, where subscript $-i$ represents player i's opponent.

3.2. Perceptual biases

The main innovation of our paper is to introduce overconfidence biases in players' beliefs. We assume that the expected distribution of one's opponent's resilience $\tilde{F}_{-i}$ is given by $F_{-i}(\hat{v} + K_i, \gamma^2/\sigma_i^2)$, where $v_{-i}^e := \hat{v} + K_i$ is the expected mean, and γ^2/σ_i^2 the expected variance, as drawn Figure 1. We are therefore considering two psychological

increases. Moreover, it is not uncommon that long wars give rise to internal disputes amongst the population, as has evidently been the case in Iraq, or Syria.

biases: an "**overestimation**" (or "**illusory superiority**") bias, measured by $K_i \geq 0$, and an "**overprecision**" bias, measured by $\sigma_i^2 \geq 1$. The first bias is the average estimation error $(v_{-i}^e - \bar{v})$ and reflects the tendency to over-estimate the opponent's type.

In our framework, the overestimation bias plays a dual role in the players' optimal dropping out decision. Overestimating the true cost of war for the opponent maps in an expectation that the opponent will be less eager to sustain a prolonged war given the lower expected gains he will obtain from winning the war. This in turn increases one's expected payoff of prolonging the war. The second role played by overestimation is that the player holding such biased beliefs anticipates lower gains from winning the war, hence reducing his incentives to prolong the war.

The second bias, overprecision, reflects the excessive certainty regarding the accuracy of own beliefs: the higher the value of σ_i, the higher the group i's overprecision bias. At the limit $(\sigma_i^2 \to +\infty)$, player i has a full self-confidence since he believes that the opponent's cost corresponds with certainty to his expected value.

Two additional points deserve attention. On the one hand, overconfidence is a rationality bias, because the expected distribution $\tilde{F}_{-i}$ differs from the true distribution F_{-i} through overconfidence parameters alone (the case without overconfidence defines a rational expectation framework, i.e., $K_i = 0$ and $\sigma_i = 1 \Rightarrow \tilde{F}_{-i} = F_{-i}$). Second, overconfidence biases give birth to an asymmetric configuration. Indeed, in the absence of biases both players share the same (and true) expected distribution

$(F_1 = F_2 = F(\hat{v}, \gamma^2))$.

Figure 1 highlights the two rationality biases. If player i is not subject to any mispreception bias, his estimation is described by the distribution $\tilde{F}^0_{-i}$, i.e. the unbiased distribution F_{-i}. If player i is subject to an overprecision bias ($\sigma_i^2 > 1$) without feelings of superiority ($K_i = 0$), his expectation of v_{-i} remains equal to $\hat{v}$, but there is a volatilities judgment "error" as depicted in the distribution $\tilde{F}^1_{-i}$. On the other hand, if player i does not suffer from overprecision ($\sigma_i^2 = 1$) but over-estimates the opponent's cost, the opponent's cost distribution is then given by $\tilde{F}^2_{-i}$, with an expected mean equal to $\hat{v} + K_i$.

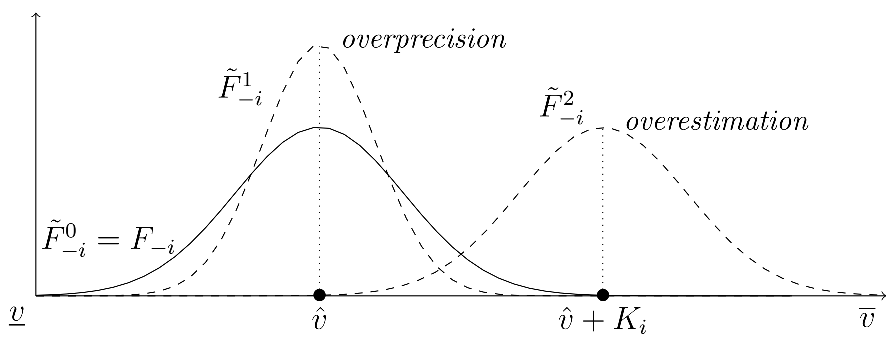

Figure 1: Overconfidence biases

3.3. Players' preferences

At each time period the players are called to make the single decision of pursuing the conflict or declaring defeat. The war therefore ends when one of the players drops out; the player that did not surrender is the *winner* labelled W and appropriates the rent of his opponent (the *loser* labelled L). If both players surrender simultaneously, they each retain control of their initial (depleted) resource, thence they obtain a depreciated value of their status quo payoff and we label that state of the world S. Thus, if player i intends to drop out at time T_i, his expected payoff at the end of war is conditional on his opponent's intentions. Assuming for simplicity that the discount rate is normalized to unit, the inter-temporal payoff of player i reads as:

$$U_i(T_i, T_{-i}, v_i) = \begin{cases} -T_i + V_i^L(T_i) & \text{if } T_i < T_{-i}, \\ -T_i + \ V_i^S(T_i) & \text{if } T_i = T_{-i}, \\ -T_{-i} + V_i^W(T_{-i}) & \text{if } T_{-i} < T_i. \end{cases} \tag{1}$$

where according to the game's description $V_i^L(T_i)$, $V_i^S(T_i)$ and $V_i^W(T_i)$ are respectively

$$V_i^L(T_i) = 0, \tag{2}$$

$$V_i^S(T_i) = R_i(T_i) = \exp(-v_i T_i), \tag{3}$$

and,

$$V_i^W(T_i) = R_i(T_i) + R_{-i}(T_i) = \exp(-v_i T_i) + \exp(-v_{-i}^e T_i). \tag{4}$$

The first line in (1) represents player i's utility if he drops out first (at time T_i), the second term if he drops out at the same time as his opponent, and the last line if the other player drops out before T_i (at time T_{-i}). Importantly, besides allowing us to describe war duration, our framework is flexible enough to capture its initiation as well since the peaceful status quo is maintained in our setting if both players simultaneously decide to drop out of the war in $t = 0$.

Let us denote by $\tilde{G}_{-i}(\cdot)$ the distribution of player i beliefs about the opponent's drop-out time (which will depend on $\tilde{F}_{-i}(v_i)$ and on his strategy, as we will see), and by $\tilde{g}_{-i}(\cdot)$ the associated density function. Player i's expected payoff is $V_i(T_i, v_i) := \mathbb{E}_i[U_i(T_i, T_{-i}, v_i)]$, hence

$$V_i(T_i, v_i) = [1 - \tilde{G}_{-i}(T_i)] \left\{-T_i + V_i^L(T_i)\right\} + \int_{x=0}^{x=T_i} \left\{-x + V_i^W(x)\right\} \tilde{g}_{-i}(x)\, dx. \quad (5)$$

As usual, a war of attrition is a noncooperative dynamic game in which players simultaneously choose at the initial instant a drop-out time to maximize their intertemporal utility according to their beliefs. In the next section we derive the game's equilibria.

4. Equilibrium

Our equilibrium concept is the bayesian equilibrium concept of Bliss and Nalebuff (1984) and Fudenberg and Tirole (1986). Player i's strategy is a couple $(s_i, T_i(v_i))$ with the second element being a (measurable) function $T_i(v_i) : [\underline{v}, \bar{v}] \to \mathbb{R}^+ \cup \{+\infty\}$, specifying for each possible value of v_i the time (possibly infinite) at which player i will surrender if the opponent is still active. Since the game features a pre-stage and a war of attrition, we proceed backwardly by first describing the equilibria of the war of attrition.

4.1. War of attrition

Consider first the following definition:

Definition 1. A pair $(T_1(v_1), T_2(v_2))$ is a bayesian equilibrium if and only if, for all $i \in \{1, 2\}$, $v_i \in [\underline{v}, \bar{v}]$, and $t \geq 0$, $V_i(T_i(v_i), v_i) \geq V_i(t, v_i)$.

War of attrition models typically exhibit multiple equilibria. One pure-strategy equilibrium is such that one player surrenders immediately at $t = 0$, and the adversary surrenders in a sufficiently distant time period so that the other player does not find it optimal to deviate by opting out at an even further time. The next proposition describes such equilibria.

Proposition 1. *There are critical times $\hat{T}_i > 0$, such that there exists an infinity of degenerate pure-strategy equilibria, where $T_i^* = 0$ and $T_{-i}^* \in [\hat{T}_i, +\infty]$, $i = \{1, 2\}$.*

Proof: See Appendix A.

These asymmetric pure strategy equilibria are efficient in the sense that no rent is "wasted" fighting. Yet, these equilibria differ from a *status quo* outcome that we explore later, since the player who surrenders immediately retains no resources and earns a zero payoff, while the winner obtains the control of both players' resources, at zero cost. In essence then, if the players' beliefs are such that both players know that one of them expects the other not to yield early, it is rational for the opponent to yield immediately.

As is standard in wars of attrition, there also exists another equilibrium in which both players drop out in a positive instant. This equilibrium is such that each player is indifferent between his pure actions, and emerges when the marginal gain and marginal cost to continue the war are balanced. The following lemma establishes some useful properties of this equilibrium with delay.

Lemma 1. *If the pair $(T_1(v_1), T_2(v_2))$ is an equilibrium with delay,*
(i) $T_i(v_i)$ is continuous and strictly decreasing on $[\underline{v}, \bar{v}]$,
(ii) $T_i(\overline{v}) = 0$.

Proof: See Appendix B.

Lemma 1 reveals that the higher-type player drops out first. Intuitively, the higher the player i's type, the higher the damages (the loss of rent) incurred in wartime, and the higher the incentive to drop out. Besides, if a player suffers the maximum possible damage ($v_i = \overline{v}$), he surrenders at the initial instant.

We can next specify the relationship between the two distributions of beliefs, namely, the distribution of the opponent's drop-out time ($\tilde{G}_{-i}(T_i)$), which is unknown, and the distribution of the opponent's type ($\tilde{F}_{-i}(v_i)$), which is fully specified. As $T_i(v_i)$ is a monotonic decreasing function, we consider $1 - \tilde{G}_{-i}(T_i(v_i)) = \tilde{F}_{-i}(v_i)$.

Let us introduce *the hazard rates* $\tilde{h}_{-i}(\cdot) := \tilde{f}_{-i}(\cdot)/\tilde{F}_{-i}(\cdot)$ denoting the probability that the opponent surrenders at the coming instant, given he has not surrender before. The following theorem characterizes the unique equilibrium with delay.

Theorem 1. *There exists a Bayesian equilibrium of the war of attrition game, which is characterized by the two following monotonically decreasing dropping-out functions $(T_1(v_1), T_2(v_2))$, where*

$$T_i'(v_i) = -\tilde{h}_{-i}(v_i)\left[e^{-v_i T_i(v_i)} + e^{-(\hat{v}+K_i)T_i(v_i)}\right], \tag{6}$$

with $T_i(\overline{v}) = 0$, for any $i = 1, 2$.

Proof: See Appendix C.

Theorem 1 states that, in equilibrium, any player i is indifferent between dropping out at t and waiting an additional increment of time before dropping out in $t + dt$. To better grasp the equilibrium condition we use the fact that $V_i^W(T_i(v_i)) - V_i^L(T_i(v_i)) = V_i^W(T_i(v_i)) = e^{-v_i T_i(v_i)} + e^{-(\hat{v}+K_i)T_i(v_i)}$ to re-write Eq. (6) as:

$$1 = \left[-\frac{1}{T_i'(v_i)} \tilde{h}_{-i}(v_i) \right] \left\{ V_i^W(T_i(v_i)) - V_i^L(T_i(v_i)) \right\}. \tag{7}$$

The left-hand side of (7) represents the cost of waiting an additional increment of time before dropping out of the war of attrition, which equals 1 given the normalization we have assumed. The right-hand side represents the expected marginal gain of waiting another instant before dropping out, which is the product of the conditional probability that the opponent drops out (the hazard rate, in brackets) and the gain in case the opponent surrenders (i.e. the net gain of victory $V_i^W(\cdot) - V_i^L(\cdot)$).[13]

To further fix the ideas, on Figure 2 we propose a visual representation of the equilibrium dropping-out plan. Consider a player of type v_i^0. The equilibrium dropping out time $T_i = T_i(v_i^0)$ is such that this player is exactly indifferent between his two pure actions, namely remaining in war, or exiting. For earlier dropping out times (e.g, $T_i = T_i^{'}$), the marginal gain exceeds the marginal cost of waiting, such that the player delays the expected instant he drops out. In contrast, at $T_i = T_i^{''}$, the marginal cost is higher than the marginal gain, and the player revises downwardly his dropping-out plan.

[13]Technically, solving the pair of differential equations $(T_1(v_1), T_2(v_2))$ in Eq. (6) yields a family of potential equilibria. To obtain a unique solution, a boundary condition is needed. Such a boundary condition is obtained by considering behavior at $v_i = \bar{v}$ (see result ii. of lemma 1).

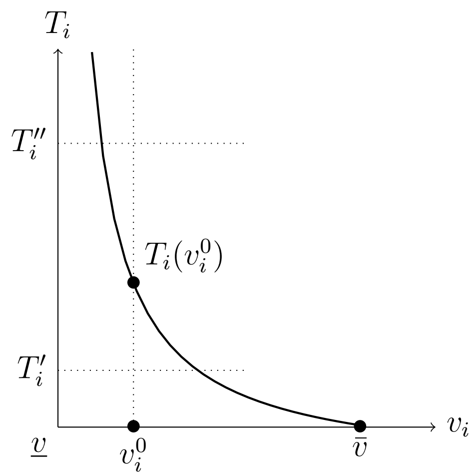

Figure 2: Optimal dropping-out strategy

Two corollaries of Theorem 1 are worth exposing at this stage.

Corollary 1. *There exists a $\bar{\bar{v}}_i \in [\underline{v}, \bar{v}]$ such that for any $v_i > \bar{\bar{v}}_i$ player i's ex-ante expected payoff in an equilibrium with strict delay is lower than its status quo payoff.*

Proof: See Appendix D.

4.2. Initiation of war

Having derived the set of equilibria for the war of attrition subgame, we now turn our attention to the declaration of war prestage. In Corollary 1 we have shown that wars of attrition that produce lower payoffs than the status quo could be observed for high enough destruction rates. Subsequently, we can straightforwardly deduce the following important corollary:

Corollary 2. *If $v_i \geq \bar{\bar{v}}_i$ for both $i = \{1, 2\}$, the status quo can be supported at equilibrium by the threat of engaging in a war of attrition with strictly positive delay (i.e. $s_i = p$, $\forall i = 1, 2$).*

This result which is straightforwardly deduced by a combination of Theorem 1 and Corollary 1 is of outmost importance since it implies that our model is able to fully describe a large spectrum of real-world IR scenarios, ranging from peaceful situations, to wars where a contender yields immediately, and long-lasting conflicts too. For this result to emerge, two features ought to be present. First, the costs of conflict need to be sufficiently high for a prolonged war to be Pareto-dominated by the status quo. Second, players must hold beliefs that give rise to the type of equilibria described in Theorem 1.

In what precedes we have derived all the game's equilibria. In the remainder of the paper, we turn our attention on the effect of overconfidence on the 's equilibria.

5. Overconfidence and war

We first consider the war of attrition equilibrium with delay, and analyse the comparative statics related to the player i's overconfidence parameters $\{K_i, \sigma_i\}$. The two following propositions determine the effect of overprecision and overestimation, respectively, on the duration of a war of attrition.

Proposition 2. *(Overprecision effect)* $\partial T_i / \partial \sigma_i \geq 0 \Leftrightarrow \partial \tilde{h}_{-i} / \partial \sigma_i \geq 0$.

Proof: See Appendix E.

This result is quite straightforward. Assume a war of attrition has been initiated. The duration of war then increases if the estimated conditional probability that the opponent drops out (the hazard rate) increases with the player's overprecision bias.

Indeed, for any estimation values, when player i over-believes that the opponent will surrender in the following instant, he is better-off continuing fighting. The effect of the over-placement in such contexts is more complex, as stated the following proposition.

Proposition 3. *(Illusory superiority effect)*

i. If $\partial\tilde{h}_{-i}/\partial K_i \leq 0$, then $\partial T_i/\partial K_i < 0$.

ii. If $\partial\tilde{h}_{-i}/\partial K_i > 0$, $\partial T_i/\partial K_i \geq 0$ if and only if

$$\left(e^{-(v_i-\hat{v}-K_i)T_i(v_i)}+1\right)\varepsilon(v_i,K_i) \geq K_i T_i(v_i), \tag{8}$$

where $\varepsilon(v_i,K_i) := \frac{\partial\tilde{h}_{-i}(v_i)}{\partial K_i}\frac{K_i}{\tilde{h}_{-i}(v_i)}$ is the elasticity of the expected hazard rate on overestimation.

Proof: See Appendix F.

Proposition 3 states that the impact of the overestimation bias (K_i) on the duration of war depends on the effect of the overestimation bias on the expected hazard rate ($\tilde{h}_{-i}(v_i)$). To clarify this mechanism, consider some arbitrary time τ. The hazard rate is the probability that the opponent will drop out at time τ, given he has not surrendered before τ. If a player's overestimation increases the estimated (unconditional) probability that the opponent will surrender at any finite time, the player holding overconfident beliefs over his opponent's distribution has now higher incentives to delay his optimal dropping-out time for two reasons: K_i increases both the estimated probability that the opponent will drop out at any earlier period to τ, but also at the precise time period τ.

If overestimation increases the estimated hazard rate, however, there may be a non-monotonic relationship between K_i and T_i (result *ii* of Proposition 3), which depends on two conflicting effects on the marginal gain of waiting another instant before dropping out of the war of attrition. Specifically, player i's overestimation increases the war duration if and only if,

$$\left(V_i^W(T_i(v_i)) - V_i^L(T_i(v_i))\right)\left[\frac{\partial \tilde{h}_{-i}(v_i)}{\partial K_i}\frac{1}{\tilde{h}_{-i}(v_i)}\right] \geq e^{-(\hat{v}+K_i)T_i(v_i)}. \tag{9}$$

The RHS of (9) represents the additional cost of a marginal increase in K_i, and this corresponds to the marginal loss to oneself of the opponent's rent's increased destruction. The LHS is the gain, which is the marginal return from victory adjusted by a factor linked to the elasticity of the estimated hazard rate (in brackets) that reflects the chances to win at the following instant. Quite intuitively, condition (9) ensures that a player's overestimation prolongs the duration of war if and only if the marginal gain exceeds the marginal cost.

To clearly highlight the potentially nonlinear effect of overestimation, we next consider the standard class of unimodal symmetric differentiable probability distributions (denoted by the set $\mathscr{D}$). With this specification, we can establish the following Proposition.

Proposition 4. *(Non-monotonic overestimation effect) Let $f \in \mathscr{D}$. For small average estimation errors, we have*

- *If $v_i < \hat{v}$, $\partial T_i / \partial K_i < 0$.*
- *If $v_i > \hat{v}$, there is a critical level $\hat{K}_i \in (0, v_i - \hat{v})$, such that*
  - *i.* $\partial T_i / \partial K_i > 0$, $\forall K_i \leq \hat{K}_i$,
  - *ii.* $\partial T_i / \partial K_i \leq 0$, $\forall K_i > \hat{K}_i$.

Proof: See Appendix G.

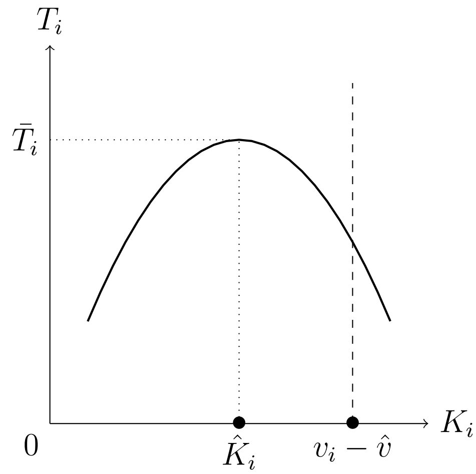

Figure 3: overestimation effect - $v_i \in [\underline{v}, \overline{v}]$

The above proposition states that if a player 's average estimation error is relatively small, two scenarios may arise. Observe first that in this case the estimated hazard rate ($\tilde{h}_{-i}$) behaves similarly to the estimated density ($\tilde{f}_{-i}$), as shown the proof in Appendix G. Then, if $v_i > \bar{v}$ the illusory superiority bias increases player i's estimated hazard rate of the opponent, while also increasing the expected damage of war on the adversary's resources. The incentives of player i to quit the war of attrition are therefore influenced in two opposing directions by his illusory superiority bias, since in delaying the optimal drop-out time player i expects to face an opponent that is expected to yield faster but who is also expected to deliver a smaller prize of war. We demonstrate that for low overestimation biases (i.e. $K_i < \hat{K}_i$) the former effect will dominate, hence pushing player i to remain active in the conflict longer. For high overestimation biases (i.e. $K_i > \hat{K}_i$) the cost of losing increases as the expected cost of war becomes higher, thereby incentivizing player i to yield earlier.

We next focus on the effect of overconfidence on the initiation of war. Corollary 2 states that, in the absence of overconfidence, a necessary condition for obtaining a peaceful outcome is to have $v_i \geq \bar{\bar{v}}_i$ for both $i \in \{1, 2\}$. Adapting the notation and defining by $\bar{\bar{v}}_i(K_i)$ the threshold value such that $\forall v_i \geq \bar{\bar{v}}_i(K_i)$ the status quo can be supported as a peaceful equilibrium, we can state the following result:

Corollary 3. $\frac{\partial T_i}{\partial K_i} \lesseqgtr 0 \Leftrightarrow \frac{\bar{\bar{v}}_i(K_i)}{\partial K_i} \gtreqless 0$

This result relates overconfidence to the set of player i's resilience parameters for which the status quo can be secured peacefully as shown in Corollary 2. In words, if overconfidence incentivizes a player to remain active longer (shorter) in a war of attrition, then this player's expected utility for a given drop-out time is necessarily higher (lower) the more overconfidence he is. This in turn implies this player's threshold value $\bar{\bar{v}}_i$ - defining the indifference between the peaceful status quo and a war of attrition - ought to be lower (higher). Thence the range of resilience parameters supporting peace at equilibrium will also be smaller (larger).

An analogous result may be stated regarding overprecision and war:

Corollary 4. $\frac{\partial T_i}{\partial \sigma_i} \geq 0 \Rightarrow \frac{\bar{\bar{v}}_i(\sigma_i)}{\partial \sigma_i} \leq 0$

This corollary reflects the content of Proposition 2 and may be interpreted along the lines of Corollary 3. Since overprecision always incentivizes a player to remain active longer in a war of attrition, we necessarily deduce that this will lead to a reduction in the range of resilience parameters supporting the peaceful status quo outcome.

6. Discussion

In what precedes we have derived a series of theoretical predictions related to the initiation and duration of wars. We have shown that wars of attrition can erupt because of information imperfections even in the absence of overconfidence, and that higher overconfidence levels may - and typically will - make such initiation more likely. Second, we have shown that overconfidence increases the duration of wars of attrition, in a possibly non-monotonic fashion.

In a number of historical contexts, overconfidence played a central role in explaining the outburst of wars, including the American Civil War, the 1904 Russo-Japanese war, or World War II. The WWI is a thoroughly studied case of leaders holding overconfident beliefs across the board (Van Evera 1999, Johnson 2004). This overconfidence can be seen in the confidence of the various actors that the war would be brief. Following the assassination of Archduke Franz Ferdinand, and in the days preceeding the imminent Austro-Hungarian war declaration on Serbia, both Berlin and Vienna were confident that (a) Russia was likely not to get involved in a military confrontation, and (b) that Russia would not constitute a major hurdle in case of armed confrontation (Johnson and Tierney 2011). Germany thought that the war would be settled in a few weeks, with Kaiser Wilhelm telling his departing troops in August 1914 “You will be home before the leaves have fallen from the trees”. In France, the military command was convinced that the French army would dominate because of the widespread enthusiasm, patriotic energy of the troops and moral superiority, and that in a few weeks time the French troops would have crossed the Rhine (Johnson 2004). Similar feelings of illusory superiority could be seen in virtually all remaining nations involved in the conflict (Johnson 2004), and yet the war proved long, and more importantly quite uncertain in its final outcome since it is the U.S. involvement in 1917 that tilted the balance in favor of the Allies. Overconfidence therefore seems to have played an important role in both the initiation and duration of World War I.

The Peloponnesian wars is another case in point that is illustrative of both the overconfidence-rooted initiation of conflict, and the failure or unwillingness of actors to update their beliefs. When Sparta decided to confront the rising city-state of Athens in 431 BC, the former had an important advantage in ground battles, while the latter excelled in navigation and sea warfare. Athens was in Sparta’s reach (250km walking), and yet the city was heavily fortified and could supply its population through its port of Piraeus if the city came under siege. Moved by their overconfidence, which was bolstered by their commonly known superiority as fighters (the mighty Spartan phalanx), “the Spartan generals remained unimaginative” (Hanson 2004:23) and applied an age-old technique of looting the Athenian hinterland so as to provoke a pitched battle outside Athens’ walls. Sparta rightly believed that if the two city-state’s armies were to meet on the battleground, the war would have been short and won by the Peloponnesians. Yet this costly strategy that mobilized thousands of Spartans away from their land and their slaves, the Helots, was stubbornly repeatedly put in action for fours consecutive years (431-427 BC) despite producing no decisive blow on their foe who remained shielded behind their city walls. It was widely recognized, however, that Athens' unmatched naval superiority would prove insufficient to annihilate their Peloponnesian foes, whose core assets lay inlands. After the initial failed raids of Sparta into Attica, Athens' homeland, it quickly became obvious to both sides that they were fighting a war of attrition that would gradually erode the belligerents' resources. From the onset, however, a rational observer would have concluded that the conflict was, by its very nature, very likely to be quite long and to be indecisive as well, thence making the peaceful resolution of the dispute the most rational approach to the issue. After all, the prediction that war would be long and indecisive can be found in Thucycides's writings, a contemporary historian providing a detailed account of the combats and strategies of the Peloponnesian wars (Thucydides 2009). Nevertheless, with the intermittence of a negotiated truce between 421 and 414 BC (truce of Nicias), the war dragged on until 404 BC, when the Spartans eventually marched into Athens and teared down its mighty walls. But two decades of highly consuming conflict had inflicted an irreversible blow to both city states alike, as well as to their allies, that would trigger the decline of Classical Greece.

The Vietnam War (1955-1975) is a third classical example whereby U.S. decision makers failed to update their beliefs on the adversary's battle-ground resilience, and got embroiled in a quagmire that cost the lives of 58,000 Americans, an estimated more than two million Vietnamese casualties, and essentially sealed a strategic defeat for the U.S. in South-East Asia. While it is difficult to integrally account for the war's length by referring to overconfidence alone, evidence suggests that throughout the war intelligence available to the US was repeatedly ignored with decision-makers retaining faulty estimates of the situation. It has indeed been argued at length that the U.S. decision-makers deliberately got involved in a very costly war of attrition despite widespread available evidence and intelligence that should have led to an early disengagement from Vietnam (e.g. Gelb and Betts 1979). Robert McNamara, who served as the secretary of Defense under the presidency of J.F. Kennedy and heavily influenced foreign policy had declared that "We have the power to knock any society out of the twentieth century". This was a correct assertion, but one disregarding the required costs the U.S. was willing to endure to achieve such a goal. The advice the White House was receiving from non-governmental sources was not to get actively involved in this conflict (Johnson 2004). Early on US president Johnson had been advised to avoid the Vietnamese quagmire. Despite the warnings received from several advisors (Johnson 2004), Kennedy who was not himself particularly enthusiastic about getting involved militarily, got heavily influenced by the hawkish voices surrounding him and nevertheless decided to send boots on the ground with the number of American soldiers reaching 16,000 shortly after Kennedy's assassination.

The best evidence of overconfidence was observed under the subsequent presidency of L.B. Johnson. By the time Johnson became president, the intelligence flowing from Vietnam had drastically increased and many sources were recommending to tone down the military intervention. For example, General Wheeler, Chairman of the Joint Chiefs of Staff, stated that total victory could only be achieved with 750,000 to a million men and up to seven years involvement - a disproportionate endeavor given the expected benefits -, while several analysts openly advised against a US involvement, notably grounding their opinions on CIA intelligence reports (Tuchman 1985). On another occasion, president Johnson was advised by retired US Army general Maxwell Taylor that "the white-faced soldier, armed, equipped, and trained as he is, not a suitable guerrilla fighter for Asian forests and jungles. The french tried to adapt their forces to this mission and failed. I doubt that US forces could do much better" (Stephenson, 2012: 362). With the years passing and the american body count increasing, the U.S. gradually intensified its bombing campaign. The U.S. logic is well summarized in General John P. McConnell's words "the military task confronting us is to make it so expensive for the North Vietnamese ... If we make it too expensive for them, they will stop" (Johnson 2004). This declaration testifies that the U.S. was consciously involved in a war of attrition, firmly believing that the opponent would yield first. The North-Vietnamese had a mutual understanding that the U.S. intervention would drag actors in a war of attrition, since their leader, Ho Chi Minh, famously declared in December 1966 that "everything depends on the Americans. If they want to make war for 20 years then we shall make war for 20 years. If they want to make peace, we shall make peace and invite them to tea afterwards". While the bombings had been raging, throughout 1966 and 1967 the CIA's estimation was that the bombing campaign would not cripple communist operations and that it was not delivering the expected results. Despite compelling evidence that the Vietnam war was not evolving as initially hoped, and in light of the important available intelligence pointing to a disengagement, U.S. decision-makers under the Johnson administration, blinded by overly optimistic beliefs in the superiority of their army, did not revise their grand strategy. Only with the election of Nixon did the White House come to terms with the reality and gradually began the retreat of the troops.

7. Conclusion

The literature on the causes and duration of wars has almost exclusively approached the question through the lenses of rational decision-makers failing to reach negotiated agreements because of informational and commitment problems. This approach is very rich and enables us to understand a vast amount of real-world confrontations. It has often been argued, however, that leaders and decision-making bodies are subject to perceptual biases which may equally contribute to deepening our understanding of why and how long nations fight each other. And yet, to date there have been very sparse efforts to study the problem from these lenses. In this paper we explore the role of overconfidence in explaining the onset and duration of conflict. More precisely, we conceptualize conflict as a war of attrition from which contenders can drop out at any point in time, therefore recognizing their defeat. The (potentially biased) decision-makers decide whether to declare war on a rival nation, and if a war is initiated, for how long to keep on fighting before accepting defeat. This seemingly simple framework delivers a host of interesting predictions.

First, we demonstrate that overconfidence is neither necessary, nor sufficient to have war. Indeed, war can arise out of asymmetric information held by rational agents failing to gauge the rival country's resilience in war. Conversely, overconfident decision-makers may decide not to declare war if they expect the benefits of peace to outmatch the gains from war. This is the case when decision-makers believe that the rival country is willing to fight long enough, and/or when the destruction generated by conflict is important. Overconfidence, however is shown to potentially reduce the scope for peaceful outcomes when biased agents are willing to remain active longer in a war of attrition, thereby being less deterred from initiating conflict in the first place. Perhaps unexpectedly, very high levels of overconfidence may well make peace easier to maintain when overconfidence maps into disproportionately high levels of damage inflicted on the opponent. In such instances, while the overconfident player is certain to achieve victory, the destruction of the loot will be such that the peaceful status quo will be preferred.

Second, our model delivers novel predictions in terms of overconfidence and the duration of war. Quite expectedly, and in line with the Peloponnesian wars, WWI, or the Vietnam War, we demonstrate that overconfidence lengthens the duration of wars. Indeed, when deciding the maximal length of time a party is willing to remain active in war before yielding to the enemy, the decision maker weighs the benefits of remaining active longer against the costs. Delaying the drop-out time increases the likelihood the rival country will yield first, but this comes at a cost of prolonging the war. Overconfident players revise their beliefs for two reasons. On the one hand they see their rivals as being more likely to yield earlier, thence incentivizing overconfident nations to remain active longer in the conflict. On the other hand, however, overconfident decision-makers also expect to inflict higher damage on their rivals and their resources, which reduces the spoils of wars and therefore tempers their desire to prolong the war. We demonstrate that for low levels of overconfidence, the former effect will most likely dominate, therefore resulting in higher war duration. If, however, the expected damage inflicted on the rival's resources is sufficiently large, the overconfident belligerent could be incentivized not to delay the war's duration as much.

Besides contributing to the sparse literature on overconfidence and war, our article also constitutes a methodological innovation to the literature on war in the continuation of Powell's (2017) model. Indeed, the use of a simple war of attrition model to describe conflict situations allows us to derive novel results. The further inclusion of extensions to such a simple setting could expand our understanding of the causes and consequences of peace and war.

[1] Antonis Adam and Petros G Sekeris. Self-containment: Achieving peace in anarchic settings. *Journal of Conflict Resolution*, 61(1):173–203, 2017.

[2] Nisha Arunatilake, Jayasuriya Sisira, and Kelegama Saman. The economic cost of the war in sri lanka. *World Development*, 29(9):1483–1500, 2001.

[3] Robert J. Aumann. Agreeing to disagree. *Annals of Statistics*, 4(6):1236–1239, 1976.

[4] Sandeep Baliga and Tomas Sjöström. Strategic ambiguity and arms proliferation. *Journal of Political Economy*, 116 (6):1023 – 1057, 2008.

[5] Muhammet A. Bas and Robert Schub. Theoretical and empirical approaches to uncertainty and conflict in international relations. *Oxford Research Encyclopedia of Politics*, forthcoming.

[6] Christopher Bliss and Barry Nalebuff. Dragon-slaying and ballroom dancing: The private supply of a public good. *Journal of Public Economics*, 25(1-2):1–12, 1984.

[7] Bruce Bueno De Mesquita. *The War Trap.* Yale University Press, 1981.

[8] Daniel L. Byman. Friend like these: Counterinsurgency and the war on terrorism. *International Security*, 31(2):79–115, 2006.

[9] Giacomo De Luca and Petros G. Sekeris. Deterrence in contests. *Economica*, 80(317):171–189, 2013.

[10] Alexandre Debs and Nuno P. Monteiro. The strategic logic of nuclear proliferation. *International Security*, 39(2):7–51, 2014.

[11] Joyce Ehrlinger, Ainsley L. Mitchum, and Carol S. Dweck. Understanding overconfidence: Theories of intelligence, preferential attention, and distorted self-assessment. *Journal of Experimental Social Psychology*, 63:94–100, 2016.

[12] James D. Fearon. Rationalist explanations for war. *International Organization*, 49 (3):379–414, 1995.

[13] James D. Fearon. Bargaining, enforcement, and international cooperation. *International Organization*, 52(2):269–305, 1998.

[14] Mark Fey and Kristopher W. Ramsey. Mutual optimism and war. *American Journal of Political Science*, 51(4):738–754, 2007.

[15] Mark Fey and Kristopher W. Ramsey. Mutual optimism in the bargaining model of war. *Unpublished manuscript, Princeton University*, 2016.

[16] Darren Filson and Suzanne Werner. A bargaining model of war and peace: Anticipating the onset, duration, and outcome of war. *American Journal of Political Science*, 46(4):819–837, 2002.

[17] Drew Fudenberg and Jean Tirole. A theory of exit in duopoly. *Econometrica*, 54:943–960, 1986.

[18] Leslie G. Gelb and Richard K. Betts. *The Irony of Vietnam: The System Worked*. The Brookings Classics, 1979.

[19] Victor Davis Hanson. *A War Like No Other: How the Athenians and Spartans Fought the Peloponnesian War*. Random House, New-York, 2005.

[20] Matthew O. Jackson and Massimo Morelli. Political bias and war. *American Economic Review*, 97(4):1353–1373, 2007.

[21] Robert Jervis. *Perception and Misperception in International Politics*. Princeton University Press, Princeton, New Jersey, 1976.

[22] Robert Jervis. War and misperception. *The Journal of Interdisciplinary History*, 18(4):675–700, 1988.

[23] Dominic D.P. Johnson. *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Harvard University Press, 2004.

[24] Dominic D.P. Johnson and Dominic Tierney. *Failing to Win: Perceptions of Victory and Defeat in International Politics*. Harvard University Press, Cambridge, MA, 2006.

[25] Dominic D.P. Johnson and Dominic Tierney. The rubicon theory of war: How the

path to conflict reaches the point of no return. *International Security*, 36(1):7–40, 2011.

[26] Daniel Kahneman and Amos Tversky. On the reality of cognitive illusions. *Psychological Review*, 103(3):582–591, 1996.

[27] Virginia S. Y. Kwan, Oliver P. John, David A. Kenny, Michael Harris Bond, and Richard W. Robins. Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. *Psychological Review*, 111:94–110, 2004.

[28] Andrew H. Kydd. *Trust and Mistrust in International Relations*. Princeton University Press, 2005.

[29] David A. Lake. Two cheers for bargaining theory: Assessing rationalist explanations of the iraq war. *International Security*, 35(3):7–52, 2010.

[30] Lawrence LeShan. *The Psychology of War: Comprehending Its Mystique and Its Madness*. Helios Press, New York, 2002.

[31] Bahar Leventoglu and Branislav L. Slantchev. The armed peace: A punctuated equilibrium theory of war. *American Journal of Political Science*, 51 (4):755–771, 2007.

[32] Bahar Leventoglu and Ahmer Tarar. Does private information lead to delay or war in crisis bargaining? *International Studies Quarterly*, 52(3):533–553, 2008.

[33] Jack S. Levy. Misperception and the causes of war: Theoretical linkages and analytical problems. *World Politics*, 36(1):76–99, 1983.

[34] Jason Lyall, Graeme Blair, and Kosuke Imai. Explaining support for combatants during wartime: A survey experiment in afghanistan. *American Political Science Review*, 107(4):679–705, 2013.

[35] Don A. Moore and Paul J. Healy. The trouble with overconfidence. *Psychological Review*, 115(2):502–517, 2008.

[36] Pietro Ortoleva and Erik Snowberg. Overconfidence in political behavior. *American Economic Review*, 105(2):504–535, 2015.

[37] Christopher Peterson. The future of optimism. *American Psychologist*, 55(1):44–55, 2000.

[38] Christopher Peterson. *A Primer in Positive Psychology*. Oxford University Press, Oxford, 2006.

[39] Robert Powell. Bargaining in the shadow of power. *Games and Economic Behavior*, 15 (2):255–289, 1996.

[40] Robert Powell. Bargaining and learning while fighting. *American Journal of Political Science*, 48(2):344–361, 2004.

[41] Robert Powell. War as a commitment problem. *International Organization*, 60:169–203, 2006.

[42] Robert Powell. Persistent fighting and shifting power. *American Journal of Political Science*, 56(3):620–637, 2012.

[43] Robert Powell. Taking sides in wars of attrition. *American Political Science Review*, 111(2):219–236, 2017.

[44] Briony D. Pulford. Overconfidence in human jugement. *PhD Thesis University of Leicester*, Department of Psychology, Leicester., 1996.

[45] Kristopher W. Ramsey. Information, uncertainty, and war. *Annual Review of Political Science*, 20:505–527, 2017.

[46] Zachary C. Shirkey. Uncertainty and war duration. *Intenational Studies Review*, 18(2):244–267, 2016.

[47] Branislav L. Slantchev. The principle of convergence in wartime negotiations. *American Political Science Review*, 97(4):621–632, 2003.

[48] Branislav L. Slantchev and Ahmer Tarar. Mutual optimism as a rationalist explanation of war. *American Journal of Political Science*, 55(1):135–148, 2011.

[49] Benjamin Smith. Oil wealth and political survival in the developing world, 1960-1999. *American Journal of Political Science*, 48 (2):232–246, 2004.

[50] Alastair Stam and Allan C. Smith. Bargaining and the nature of war. *Journal of Conflict Resolution*, 48(6):783–813, 2004.

[51] Michael Stephenson. *The Last Full Measure: How Soldiers Die in Battle*. Crown Publishing, New York, 2012.

[52] John G. Stoessinger. *Why Nations Go To War?* St Martin's, New-York, 2011.

[53] Patricia L. Sullivan. War aims and war outcomes: Why powerful states lose limited wars. *Journal of Conflict Resolution*, 51(3):496–524, 2007.

[54] Shelley E. Taylor and Jonathon D. Brown. Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103:193–210, 1988.

[55] Shelley E. Taylor and Jonathon D. Brown. Positive illusions and well-being revisited: Separating fact from fiction. *Psychological Bulletin*, 116:21–27, 1994.

[56] Thucydides. *The Peloponnesian War*. Oxford World's Classics. A new translation by Martin Hammond, 2009.

[57] Barbara W. Truchman. *The March of Folly: From Troy to Vietnam*. Abacus, 1985.

[58] Stephen Van Evera. *Causes of War: Power and the Roots of Conflict*. Cornell University Press, 1999.

[59] R. Harrison Wagner. Bargaining and war. *American Journal of Political Science*, 44(3):469–484, 2000.

[60] Donald Wittman. How a war ends: A rational model approach. *Journal of Conflict Resolution*, 23(4):743–763, 1979.

Appendix A. The degenerate pure-strategy equilibrium

By Eqs. (1), (2), (3), and (4), the payoff of player 1 and 2 are, respectively

$$U_1(T_1, T_2, v_1) = \begin{cases} -T_1 & \text{if } T_1 < T_2, \\ -T_1 + e^{-v_1 T_1} & \text{if } T_1 = T_2, \\ -T_2 + e^{-v_1 T_2} + e^{-(\hat{v}+K_1)T_2} := h_1(T_2) & \text{if } T_2 < T_1. \end{cases} \tag{A.1}$$

$$U_2(T_2, T_1, v_2) = \begin{cases} -T_2 & \text{if } T_2 < T_1, \\ -T_2 + e^{-v_2 T_2} & \text{if } T_1 = T_2, \\ -T_1 + e^{-v_2 T_1} + e^{-(\hat{v}+K_2)T_1} =: h_2(T_1) & \text{if } T_1 < T_2. \end{cases} \tag{A.2}$$

We are interested in the best response correspondences. Let us compute the player 1's best response function. Observe first that $h_1 \in C^\infty(\mathbb{R}_+)$, h_1 is strictly decreasing in T_2, $h_1(0) = 2 > 0$, and $h_1(+\infty) = -\infty < 0$. By the Intermediate Value Theorem we know that there exists a unique critical value $\hat{T}_1$ (that depends on v_1 and K_1), such that: $T_2 < \hat{T}_1 \Leftrightarrow h_1(T_2) > 0$ *(case a)*, and $T_2 > \hat{T}_1 \Leftrightarrow h_1(T_2) < 0$ *(case b)*.

Case a. $T_2 < \hat{T}_1$. In this case the payoff function is depicted in Figure A.4.

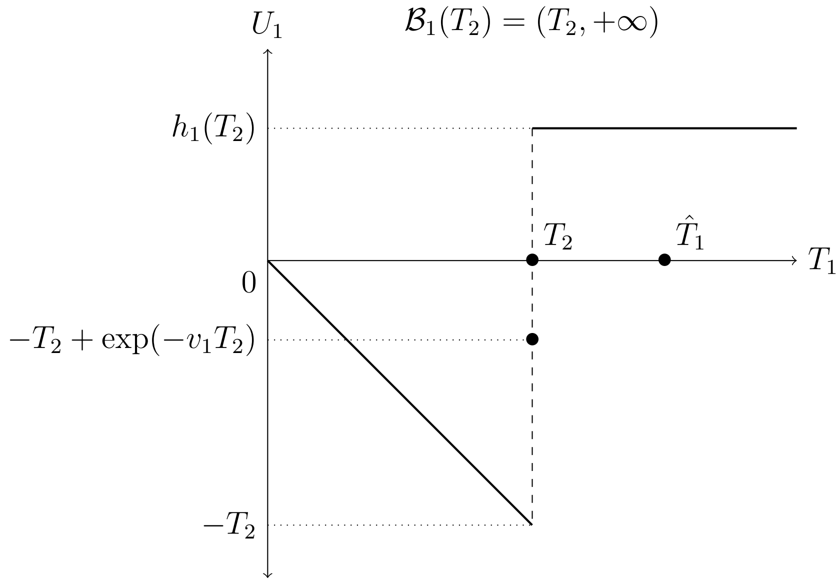

Figure A.4: Case (a) $T_2 < \hat{T}_1$

Case b. $T_2 > \hat{T}_1$. In this case the payoff function is depicted in Figure A.5.

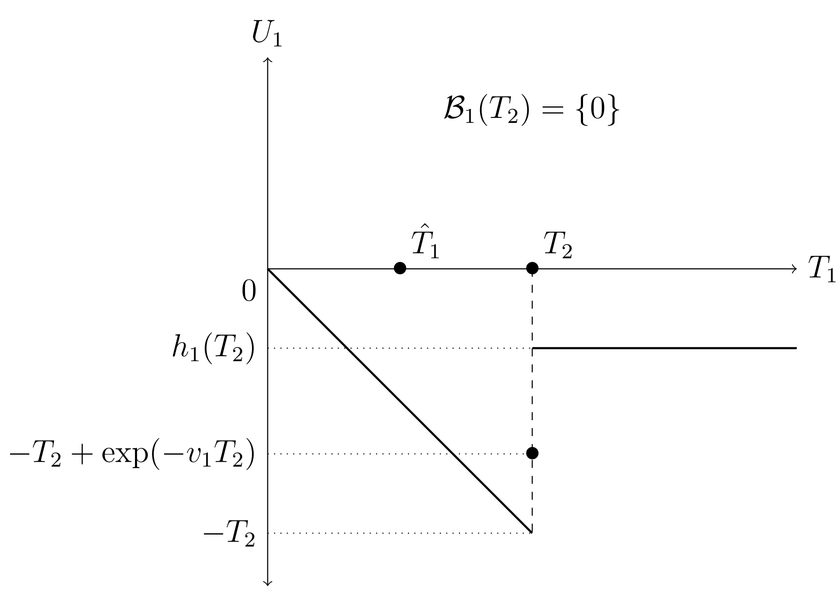

Figure A.5: Case (b) $T_2 > \hat{T}_1$

Player 1's best response correspondence, $\mathcal{B}_1(T_2)$ is thus given by

$$\mathcal{B}_1(T_2) = \begin{cases} (T_2, +\infty) & \text{if } T_2 < \hat{T}_1, \\ \{0\} \cup (T_2, +\infty) & \text{if } T_2 = \hat{T}_1, \\ \{0\} & \text{if } T_2 > \hat{T}_1. \end{cases}$$

Similarly, player 2's best response correspondence $\mathcal{B}_2(T_1)$ is

$$\mathcal{B}_2(T_1) = \begin{cases} (T_1, +\infty) & \text{if } T_1 < \hat{T}_2, \\ \{0\} \cup (T_1, +\infty) & \text{if } T_1 = \hat{T}_2, \\ \{0\} & \text{if } T_1 > \hat{T}_2, \end{cases}$$

Consequently, combining $\mathcal{B}_1(T_2)$ and $\mathcal{B}_2(T_1)$, we deduce that any pair (T_i^*, T_{-i}^*), $i = \{1, 2\}$, with $T_i^* = 0$ and $T_{-i}^* \in [\hat{T}_i, +\infty]$ is an equilibrium.

Appendix B. Proof of Lemma 1

We want to inspect the sign of $\partial T_i(v_i)/\partial v_i$. We thus proceed as follows. Using Eqs. (2), (3) and (4), the expected payoff (5) becomes

$$V_i(T_i, v_i) = -T_i[1 - \tilde{G}_{-i}(T_i)] + \int_{x=0}^{x=T_i} \left\{-x + e^{-v_i x} + e^{-(\hat{v}+K_i)x}\right\} \tilde{g}_{-i}(x)\, dx. \quad \text{(B.1)}$$

By differentiating with respect to T_i, we obtain

$$\frac{\partial V_i(T_i, v_i)}{\partial T_i} = -[1 - \tilde{G}_{-i}(T_i)] + \tilde{g}_{-i}(T_i) \left\{e^{-v_i T_i} + e^{-(\hat{v}+K_i)T_i}\right\} =: \psi(T_i, v_i). \quad \text{(B.2)}$$

Let us suppose that there is an interior optimal dropping-out time T_i, a feature that we later prove in Appendix C. In that case, for any $v_i \in [\underline{v}, \overline{v}]$, T_i must satisfy the first-order condition which reads as: $\psi(T_i, v_i) = 0$. To obtain the sign of $\partial T_i(v_i)/\partial v_i$ we thus apply the Implicit Function Theorem to $\psi(T_i, v_i) = 0$, and deduce that

$$\frac{\partial T_i}{\partial v_i} = -\frac{\partial \psi(T_i, v_i)/\partial v_i}{\partial \psi(T_i, v_i)/\partial T_i}. \tag{B.3}$$

Since $\psi(T_i, v_i) = 0$ characterizes an interior solution to the problem, the second-order condition is also verified, and thus $\partial \psi(T_i, v_i)/\partial T_i < 0$. We therefore deduce that $\operatorname{sgn}\{\partial T_i(v_i)/\partial v_i\} = \operatorname{sgn}\{\partial \psi(T_i, v_i)/\partial v_i\}$, which is given by

$$\frac{\partial \psi(T_i, v_i)}{\partial v_i} = \frac{\partial^2 V_i}{\partial T_i\, \partial v_i} = -T_i \tilde{g}_{-i}(T_i) e^{-v_i T_i} < 0. \tag{B.4}$$

Thus, the optimal dropping-out plan $T_i(\cdot)$ is a monotonic decreasing function in v_i.

□

Appendix C. Equilibrium with delay

We derive the bayesian equilibrium using a three-steps proof. We compute the first-order condition (step 1), establish the existence (step 2), and then prove that the second-order condition is verified (step 3).

i. First-order condition. According to Definition 1, the couple $(T_1(v_1), T_2(v_2))$ is a Bayesian equilibrium if each function $T_i(v_i)$ maximizes player i's expected payoff, for any $i \in \{1, 2\}$. As T_i is monotonic with respect to v_i (see Lemma 1), choosing a drop-out time T_i as in the previous section is equivalent to choosing a value $\check{v}_i$, and dropping out at time $T_i = T_i(\check{v}_i)$. Using this alternative approach and re-writing the drop out time as a function of $\check{v}_i$ we thus have

$$\frac{\partial V_i(T_i(\check{v}_i), v_i)}{\partial \hat{v}_i} = T_i'(\check{v}_i)\frac{\partial V_i(T_i(\check{v}_i), v_i)}{\partial T_i}. \tag{C.1}$$

Using (B.1), the first-order condition is, after the change in variables[14]

$$\frac{\partial V_i(T_i(\check{v}_i), v_i)}{\partial \hat{v}_i} = -T_i'(\check{v}_i)\tilde{F}_{-i}(\check{v}_i) - \tilde{f}_{-i}(\check{v}_i)\left\{e^{-v_iT_i(\check{v}_i)} + e^{-(\hat{v}+K_i)T_i(\check{v}_i)}\right\}. \tag{C.2}$$

Since $T_i(v_i)$ is the optimal dropping-out time for the player i with cost v_i, we have $\check{v}_i = v_i$ when $\check{v}_i$ is chosen optimally. Namely, the FOC (C.2) evaluated at $\hat{v}_i = v_i$ implies

$$T_i'(v_i) = \left[-\frac{\tilde{f}_{-i}(v_i)}{\tilde{F}_{-i}(v_i)}\right]\left(e^{-v_iT_i(v_i)} + e^{-(\hat{v}+K_i)T_i(v_i)}\right). \tag{C.3}$$

Using $\tilde{h}_{-i}(\cdot) := \tilde{f}_{-i}(\cdot)/\tilde{F}_{-i}(\cdot)$, Eq. (6) in the main text immediately follows.

ii. Existence. Eq. (C.3) corresponds to the initial value problem $T_i' = \phi(v_i, T_i)$, with

$$\phi(v_i, T_i) = -\tilde{h}_{-i}(v_i)\left(e^{-v_iT_i(v_i)} + e^{-(\hat{v}+K_i)T_i(v_i)}\right), \text{ and } T_i(\overline{v}) = 0,$$

Let $\epsilon > 0$ an arbitrary small scalar, and let $\mathcal{V} = [\underline{v} + \epsilon, \overline{v}]$ a closed interval. Consequently, $\tilde{h}_{-i}(v_i) > 0$ for any $v_i \in \mathcal{V}$, hence functions $\phi(\cdot,\cdot)$ and $\partial_{T_i}\phi(\cdot,\cdot)$ are continuous on $\mathcal{V} \times [0, +\infty)$. Thus, according to the Picard's theorem, there exists at least one solution to the initial value problem.

iii. Second-order condition. Substituting $T_i'(v_i)$ evaluated at $\check{v}_i$ from (C.3) into

[14] We use $\tilde{G}_{-i}(T_i(\check{v}_i)) = 1 - \tilde{F}_{-i}(\check{v}_i)$, hence; $T_i'(\check{v}_i)\tilde{g}_{-i}(T_i(\check{v}_i)) = -\tilde{f}_{-i}(\check{v}_i)$.

(C.2), we obtain

$$\frac{\partial V_i(T_i(\check{v}_i), v_i)}{\partial \hat{v}_i} = \tilde{f}_{-i}(\check{v}_i)\left[e^{-\check{v}_i T_i(\check{v}_i)} + e^{-(\hat{v}+K_i)T_i(\check{v}_i)} - e^{-v_i T_i(\check{v}_i)} - e^{-(\hat{v}+K_i)T_i(\check{v}_i)}\right].$$

Differentiating with respect to $\check{v}_i$ and considering $v_i = \check{v}_i$, it follows that

$$\left.\frac{\partial^2 V_i(T_i(\check{v}_i), v_i)}{\partial \check{v}_i^2}\right|_{v_i=\check{v}_i} = -\tilde{f}_{-i}(v_i)T_i(v_i)e^{-v_i T_i(v_i)} < 0,$$

hence, the second-order condition is verified.

Appendix D. Proof of Corollary 2

Let us consider the equilibrium with delay. We first compute

$$\frac{\partial V_i(T_i(v_i), v_i)}{\partial v_i} = T_i'(v_i)\frac{\partial V_i(T_i(v_i), v_i)}{\partial T_i} + \frac{\partial V_i(T_i(v_i), v_i)}{\partial v_i}$$

At equilibrium, we have (i) $\partial V_i/\partial T_i = 0$, and (ii) $\partial V_i/\partial v_i < 0$, as demonstrated in Appendix B. Consequently, the utility of player i is decreasing in v_i. Using (B.1), as $T_i(\overline{v}) = 0$, we have $V_i(T_i(\overline{v}), \overline{v}) = 0$. Thus, according to the Intermediate Value Theorem, if $V_i(T(\underline{v}), \underline{v}) > R_i(0) := 1$, there exists a unique value $\bar{\bar{v}}_i \in]\underline{v}, \overline{v}]$,[15] such that $V_i(T_i(v_i), v_i) > 1$ for $v_i \in (\underline{v}, \bar{\bar{v}}_i)$; and $V_i(T_i(v_i), v_i) < 1$ for $v_i \in (\bar{\bar{v}}_i, \overline{v})$.

Appendix E. Proof of Proposition 2

The optimal function drop-out function $T_i(\cdot)$ is implicitly defined by the FOC (C.2), under the boundary condition $T_i(\overline{v}) = 0$. From appendix C, the second-order

[15] If $V_i(T(\underline{v}), \underline{v}) < 1$, we have $\bar{\bar{v}}_i = \underline{v}$.

condition is satisfied, namely $\partial^2 V_i/\partial \check{v}_i^2 < 0$. Consequently, according to the Implicit Function Theorem, we have

$$\frac{\partial \check{v}_i}{\partial \sigma_i} = -\frac{\partial V_i^2/\partial \check{v}_i \partial \sigma_i}{\partial V_i^2/\partial \check{v}_i^2},$$

hence $\text{sgn}(\partial \check{v}_i/\partial \sigma_i) = \text{sgn}(\partial V_i^2/\partial \check{v}_i \partial \sigma_i)$. In addition, as $T_i(\check{v}_i)$ is monotonically decreasing in $\check{v}_i$ (Lemma 1), $T_i(\check{v}_i)$ increases in σ_i if and only if $\partial V_i^2/\partial \check{v}_i \partial \sigma_i < 0$.

By differentiating (C.2) with respect to σ_i, we obtain

$$\frac{\partial^2 V_i}{\partial \check{v}_i \partial \sigma_i} = -T_i'(\check{v}_i)\frac{\partial \tilde{F}_{-i}(\check{v}_i)}{\partial \sigma_i} - \frac{\partial \tilde{f}_{-i}(\check{v}_i)}{\partial \sigma_i}\left\{e^{-v_i T_i(\check{v}_i)} + e^{-(\hat{v}+K_i)T_i(\check{v}_i)}\right\}. \tag{E.1}$$

Therefore, using Eq. (C.3), we find

$$\frac{\partial^2 V_i}{\partial \check{v}_i \partial \sigma_i} \leq 0 \Leftrightarrow -\frac{\partial \tilde{F}_{-i}(\check{v}_i)}{\partial \sigma_i}\tilde{f}_{-i}(\check{v}_i) + \frac{\partial \tilde{f}_{-i}(\check{v}_i)}{\partial \sigma_i}\tilde{F}_{-i}(\check{v}_i) \leq 0 \Leftrightarrow -\frac{1}{(\tilde{F}_{-i}(\check{v}_i))^2}\frac{\partial \tilde{h}_{-i}(\check{v}_i)}{\partial \sigma_i} \leq 0.$$

Thus, as $\check{v}_i = v_i$ at equilibrium, $T(v_i)$ increases in σ_i iff $\partial \tilde{h}_{-i}(v_i)/\partial \sigma_i \geq 0$, $\forall v_i \in [\underline{v}, \overline{v}]$.

□

Appendix F. Proof of Proposition 3

This proof follows the lines of the proof of Proposition 2. We have that $T_i(\check{v}_i)$ decreases in K_i iff $\partial V_i^2/(\partial \check{v}_i \partial K_i) \geq 0$. Therefore, by differentiating (C.2) with respect to K_i, we obtain

$$\frac{\partial V_i}{\partial \check{v}_i \partial K_i} = -T_i'(\check{v}_i)\frac{\partial \tilde{F}_{-i}(\check{v}_i)}{\partial K_i} - \frac{\partial \tilde{f}_{-i}(\check{v}_i)}{\partial K_i}\left\{e^{-v_i T_i(\check{v}_i)} + e^{-(\hat{v}+K_i)T_i(\check{v}_i)}\right\} + T_i(\check{v}_i)\tilde{f}_{-i}(\check{v}_i)e^{-(\hat{v}+K_i)T_i(\check{v}_i)}. \tag{F.1}$$

Thus, using Eq. (C.3), we find

$$\frac{\partial^2 V_i}{\partial \check{v}_i \partial K_i} \geq 0 \Leftrightarrow -\left(e^{-v_i T_i(\check{v}_i)} + e^{-(\hat{v}+K_i)T_i(\check{v}_i)}\right) \frac{\partial \tilde{h}_{-i}(\check{v}_i)}{\partial K_i} + T_i(\check{v}_i)\tilde{h}_{-i}(\check{v}_i)e^{-(\hat{v}+K_i)T_i(\check{v}_i)} \geq 0. \tag{F.2}$$

We first observe that $\partial \tilde{h}_{-i}/\partial K_i < 0 \Rightarrow \partial T_i/\partial K_i \leq 0$. In addition, $\partial T_i/\partial K_i \leq 0$ if and only if Eq. (F.2) evaluated at $v_i = \hat{v}_i$ is satisfied. $\square$

Appendix G. Proof of Proposition 4

As $\tilde{h}_{-i}(\cdot) = \tilde{f}_{-i}(\cdot)/\tilde{F}_{-i}(\cdot)$, the estimated hazard rate can be written as

$$\tilde{h}_{-i}(v_i) = \frac{d}{dv_i} \ln(\tilde{F}_{-i}(v_i)). \tag{G.1}$$

By considering an unimodal symmetric differentiable distribution, the mode just equals the mean $(\hat{v}+K_i)$, hence $\tilde{F}_{-i}(\hat{v}+K_i) = 1/2$. By linearizing in the neighborhood of the mean $(\hat{v}+K_i)$, we obtain

$$\ln(\tilde{F}_{-i}(v_i)) \approx \ln\left(1 + [2\tilde{F}_{-i}(v_i) - 1]\right) - \ln(2) = 2\tilde{F}_{-i}(v_i) - 1 - \ln(2).$$

Using (G.1), the estimated hazard rate in the neighborhood of $(\hat{v}+K_i)$ is $\tilde{h}_{-i}(v_i) \approx 2\tilde{f}_{-i}(v_i)$. As $\tilde{f}_{-i}$ is unimodal and symmetric, it follows that, for values of v_i close to $\hat{v}+K_i$,

$$\begin{cases} \partial \tilde{h}_{-i}(v_i)/\partial K_i \geq 0 \Leftrightarrow \partial \tilde{f}_{-i}(v_i)/\partial K_i \geq 0, & \text{if } v_i - \hat{v} \geq K_i, \\ \partial \tilde{h}_{-i}(v_i)/\partial K_i < 0 \Leftrightarrow \partial \tilde{f}_{-i}(v_i)/\partial K_i < 0, & \text{if } v_i - \hat{v} < K_i. \end{cases}$$

From Proposition 3, if $K_i > v_i - \bar{v} \Rightarrow \partial\tilde{h}_{-i}(v_i)/\partial K_i < 0 \Rightarrow \partial T_i(v_i)/\partial K_i < 0$. If $K_i \leq v_i - \bar{v}$, in contrast, we have $\partial T_i(v_i)/\partial K_i \geq 0 \Leftrightarrow \phi(K_i) \leq 0$. The function $\phi(\cdot)$ derives from the left-hand side of Eq. (F.2) by considering K_i close enough to $v_i - \bar{v}$, namely

$$\phi(K_i) := -2\left[\frac{\partial f_{-i}(v_i)}{\partial K_i}\frac{1}{\tilde{f}_{-i}(v_i)}\right] + T_i(v_i).$$

Thus, ϕ is a differentiable function on $(0, v_i - \bar{v})$, with $\phi(v_i - \bar{v}) = T_i(v_i - \bar{v}) > 0$, and

$$\phi'(K_i) = -\frac{2}{\tilde{f}_{-i}(v_i)}\left[\frac{\partial^2\tilde{f}_{-i}(v_i)}{\partial K_i^2} - \frac{\partial\tilde{f}_{-i}(v_i)}{\partial K_i}\frac{1}{\tilde{f}_{-i}(v_i)}\right] + \frac{\partial T_i(v_i)}{\partial K_i},$$

hence;

$$\phi'(K_i) > 0 \Leftrightarrow -\frac{2}{\tilde{f}_{-i}(v_i)}\left[\frac{\partial^2\tilde{f}_{-i}(v_i)}{\partial K_i^2} - \frac{\partial\tilde{f}_{-i}(v_i)}{\partial K_i}\left(\frac{1}{\tilde{f}_{-i}(v_i)} + 1\right)\right] + T_i(v_i) \geq 0,$$

As $\tilde{f}_{-i}(\cdot)$ is unimodal and symmetric, we have $\max\tilde{f}_{-i}(v_i) = \tilde{f}_{-i}(\bar{v} + K_i)$. Thus, if v_i is close to $\bar{v} + K_i$, $\partial^2\tilde{f}_{-i}(v_i)/\partial K_i^2 < 0$ and $\partial\tilde{f}_{-i}(v_i)/\partial K_i \approx 0$, and it follows that $\phi'(K_i) > 0$.

Consequently, there is a critical value $\hat{K}_i \in (0, v_i - \bar{v})$, such that: $\phi(K_i) < 0$, i.e. $\partial T_i(v_i)/\partial K_i > 0$, if $K_i < \hat{K}_i$; and $\phi(K_i) > 0$, i.e. $\partial T_i(v_i)/\partial K_i < 0$, if $K_i > \hat{K}_i$. $\square$