**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

How does size matter for military success? Evidence from virtual worlds

Mildenberger, Carl David ; Pietri, Antoine

DOI: <https://doi.org/10.1016/j.jebo.2018.08.012>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-191831>

Journal Article

Published Version

The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Mildenberger, Carl David; Pietri, Antoine (2018). How does size matter for military success? Evidence from virtual worlds. *Journal of Economic Behavior & Organization*, 154:137-155.

DOI: <https://doi.org/10.1016/j.jebo.2018.08.012>

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo How does size matter for military success? Evidence from virtual worlds[☆]Carl David Mildenberger^a, Antoine Pietri^{b,*}^a University of St. Gallen, Department of Philosophy, Unterer Graben 21, St. Gallen 9000, Switzerland^b Institute for Strategic Research (IRSEM, Paris), École Militaire, 1 Place Joffre – Case 38, Paris SP 07 75 700, France

ARTICLE INFO

Article history:

Received 24 January 2018

Revised 5 August 2018

Accepted 11 August 2018

Available online 28 August 2018

JEL classification:

D74

H56

C57

Keywords:

Contest success functions

Military victory

Virtual worlds

ABSTRACT

This paper investigates how differences in military spending translate into probability of victory in a conflict. To do so the paper empirically estimates and compares the four main forms of contest success functions – the *Tullock*, the *logit*, the *difference*, and the *relative difference form*. In order to circumvent measurement issues and endogeneity biases associated with historical battle-related data, we advocate the use of virtual worlds. Data from virtual worlds is an innovative and promising tool in conflict research. It allows conflict analyses based on rich and objective empirical evidence on economic behavior in a warfare context; something which is difficult to achieve in real world settings. Thanks to collaboration with the developer of a virtual world, we are able to construct an original database of 19,229 battles that occurred during January 2011. The results show that the relative difference contest success function as proposed by Beviá and Corchón (2015) outperforms other existing forms of contest success functions. Thus, the decisive factor to predict the outcome of a battle is the relative difference in forces devoted by rivals. *Relative size matters*.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Since Pareto's observation that "[t]he efforts of men are utilized in two different ways: they are directed to the production or transformation of economic goods, or else to the appropriation of goods produced by others" (Pareto, 1971, p. 341, Section 17), appropriation and conflict gradually made their way into economic analysis. The most widely used economic models of violent appropriation are "guns versus butter" (hereafter 'GVB') models (seminal contributions are Hirshleifer, 1988; Hirshleifer, 1989).¹ In GVB models, rival agents compete over a prize. To increase the probability of success, each agent can devote resources to buy 'guns' (i.e. exercise effort to fight, grab, or extort). As guns cannot be used to create wealth, there is a trade-off between producing goods ('butter') or appropriating them from others.

[☆] This paper hugely benefitted from insightful remarks and suggestions from Fancis Bloch, Raul Caruso, Luis Corchón, Bertrand Crettez, José de Sousa, Kai Konrad, Claude Ménard, Petros Sekeris, Stergios Skaperdas and Mehrdad Vahabi. We would also like to thank two anonymous reviewers for JEBO and the editor Daniel Houser for their very valuable comments. Last but not least, we would like to express our gratitude to Thomas B. for his help with data processing. All mistakes remain ours.

* Corresponding author.

E-mail addresses: carldavid.mildenberger@unisg.ch (C.D. Mildenberger), antoine.pietri@irsem.fr (A. Pietri).

¹ These can easily be integrated into numerous economics frameworks including Edgeworth box analyses (Anderton and Carter, 2008; 2009), international trade (Garfinkel et al., 2012; 2015), growth models (Aizenman and Glick, 2006) or social network (Franke and Öztürk, 2015). The GVB framework constitutes a bridge between conflict and economics and allows a better understanding of economic activities.

GVB models depict conflict as a *costly lottery* in the sense that the probability of success is defined by a mathematical function depending on the fighting efforts devoted by each agent. A fundamental research question is to determine the form of the function explaining how increased fighting efforts are translated into an increased probability of winning, the so-called *Contest Success Function* (hereafter ‘CSF’).² Four forms received special attention in the literature. The *Tullock* form (Tullock, 1980) considers that the probability of winning is a function of the relative number of guns deployed on the battlefield by each side. The *logit* and *probit* forms (respectively proposed by Hirshleifer, 1988; Hirshleifer, 1989 and Baik, 1998) assume that victory depends on the absolute difference in guns involved in the battlefield. Lastly, the *Relative Difference CSF* (RDCSF), introduced by Beviá and Corchón (2015), considers that the key factor is the difference in fighting efforts scaled by the size of the conflict.

As pointed out by Hwang (2012), economic predictions and policy advice drastically hinge on the form of the CSF used. Consequently, comparing the different forms of CSFs is of critical importance for theoretical frameworks – in particular those dealing with countries’ optimal resource allocation. However, few empirical comparisons have been carried out as regards human warfare (e.g. Lotzin, 2011 or Hwang, 2012).³ power to predict. His conclusion is in favor of the *probit* form. Yildizparlak (2017) uses European soccer leagues’ data to estimate and compare CSFs accounting for the probability of a draw. The paper shows that the difference of efforts are more predictive than the aggregate level of effort. This is partly due to the fact that, for practical and ethical reasons, it is hard to get a hold of good, complete, and reliable data on human battles. Measurement and endogeneity biases further complicate the empirical estimation and comparison of CSFs (Jia et al., 2013). In light of these known issues, this article proceeds to use data taken from a virtual world to shed new light on the nature of the link between resources devoted to guns and the probability of success.

Virtual worlds are computer-created environments that visually mimic complex physical spaces, where people can interact with each other and with virtual objects in multiple ways (Bainbridge, 2007). They give rise to sophisticated political systems governed by distinct sets of institutions and well-developed economies (e.g. Castronova, 2001; Castronova, 2008; Lastowka and Hunter, 2004; Lessig, 1999; Ludlow, 2001 or Mildenberger, 2013a). Virtual worlds constitute a promising source of data as they provide a large amount of extremely precise data. Fine-grained analyses of battlefield behavior are possible as everything a user does within these worlds is registered on the servers. Thanks to a collaboration with the developer of a virtual world, we collected data on battles comprising more than 300,000 casualties in an online game called *EVE Online* (hereafter ‘EVE’).⁴ EVE is a real-time massively multiplayer online game which closely fits a GVB framework. That is, EVE is neither an arcade game like *Space Invaders* nor a real-time strategy game like the *Command and Conquer* series, which both possess CSFs which are predetermined by the games’ developers’. In EVE, you cannot let the computer fight for you – on “autopilot”, so to speak – to let the game’s “battle algorithm” automatically determine the result, but it is real people’s software-mediated decisions and actions alone which decide the battle. The world of EVE is a virtual anarchic State of nature, in which the around 400,000 players of this game constantly face the paradigmatic choice of either making a living by producing things or by appropriating what others produced.

We built an original database made up of 19,229 virtual battles between two ‘teams’ of players. Among the four forms of CSFs cited above, we find robust evidence advocating the RDCSF (Beviá and Corchón, 2015). More generally we find that measurements taking into account relative difference of fighting efforts outperform measurements based on absolute difference or ratio measurements. In other words, relative size matters for military success. Because the choice of a CSF has dramatic effects on the predictions of a model of conflict, this finding is important. However, the article’s contribution is also methodological. The comparison of CSFs based on virtual world data serves to exemplify the power and usefulness of the new approach of virtual-worlds research.

The article proceeds as follows. In the second section, we briefly introduce the four main forms of CSFs, as well as the results of the comparisons of CSFs by Lotzin (2011) and Hwang (2012). We also present the known data issues induced by the use of historical data sets. In the third section, we describe the economic, political, and military environment of EVE. Our fourth section describes our data, as well as the advantages and limitations of virtual-worlds research. Our identification strategy for battles, estimations and comparisons are presented in the fifth section. Section 6 presents our results, Section 7 concludes.

2. Theoretical and empirical basics of CSFs

CSFs map how fighting efforts devoted by each contestant are translated into an increased or lowered probability of success in conflict. The literature identifies different types of CSFs.

² CSFs are also used in several other different contexts like litigation, technology races, lobbying, or sports. For an overview see for example Konrad (2009), Van Long (2013) or Corchón and Serena (2017). Surveys focusing on military contests can be found in Garfinkel and Skaperdas (2007) and Jia and Skaperdas (2012).

³ By contrast, there exists a vast empirical literature on sports contests dealing with the estimation of the link between efforts devoted by sportsmen and their probability of success (e.g. golf (Ehrenberg and Bognanno, 1990), soccer (Maloney and McCormick, 2000) or tennis (Sunde, 2009)). Yet again, there are few papers addressing the comparison of the predictive performance of different CSFs for sports contests. Peeters (2011) collects data from the American major sports leagues and finds that the *Tullock* form provides better predictions than the *logit* one for basketball (NBA) and football (NFL). However, the difference of fit is not significant for hockey (NFL) and baseball (MLB). Jia (2008) focuses on the prediction of NBA game results in the 2004–2005 season. He relies on a Bayesian model comparison method to compare *Tullock*, *logit* and *probit* CSFs.

⁴ In recent years, video games have been widely used in the context of conflict theory: see for example Mildenberger (2015) or Robinson (2016).

2.1. Four forms of CSFs

Consider two teams facing each other on a battlefield. Fighting efforts, G_i , correspond to the amount of resources committed to the fight by team i , $i = 1, 2$. Let $p_i(G_i, G_j)$ be the probability of success of team i embodied by a CSF, such that $p_1 + p_2 = 1$, $\partial p_i / \partial G_i > 0$ and $\partial p_i / \partial G_j < 0$, for $j = 1, 2$, $i \neq j$ (Skaperdas, 1996). In this article, we use three types of measurement to capture the advantage of Team 1: the ratio of fighting efforts ($x_r = \ln(G_1/G_2)$), their absolute difference ($x_d = G_1 - G_2$) and their relative difference ($x_{rd} = (G_1 - G_2)/(G_1 + G_2)$). If, x_k , $k = \{r, d, rd\}$, is strictly positive this means that Team 1 commits more resources to the battlefield than Team 2 ($G_1 > G_2$). On the contrary, when $G_1 < G_2$, Team 1 has a disadvantage on the battlefield ($x_k < 0$).

For all battles involving positive fighting efforts of each team ($G_i > 0$, $\forall i$), one can categorize most existing forms of CSFs in one of three families of functions:⁵

$$\psi_1(x_k) = \frac{1}{1 + \exp(-mx_k)} \quad (1)$$

$$\psi_2(x_k) = \int_{-\infty}^{x_k} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{mx_k^2}{2}\right) dx_k \quad (2)$$

$$\psi_3(x_k) = \min \left\{ \max \left\{ \frac{1}{2} + mx_k, 0 \right\}, 1 \right\} \quad (3)$$

where x_k is the measurement of the difference of fighting efforts, and $m \geq 0$ corresponds to the decisiveness parameter. The value of this parameter determines the degree “to which a side’s greater fighting effort translates into enhanced battle success” (Hirshleifer, 2000, p. 776). In other words, m captures the level of noise in the conflict: for $m = 0$ the winner is taken as random and for $m \rightarrow \infty$ the conflict is fully discriminatory in the sense that the team investing more fighting efforts in the battlefield wins with probability one.⁶ With these elements we are able to represent the four main forms of CSF existing in the literature.

A first canonical CSF is the *Tullock* form, initially designed by Tullock (1980). It is also known as the *ratio* form. It describes the probability of victory on the battlefield of Team 1 as a function of the ratio of fighting efforts devoted by each rival:

$$p_1 = \begin{cases} \psi_1(x_r) = \frac{1}{1 + \left(\frac{G_2}{G_1}\right)^m} & \text{if } G_1 + G_2 > 0, \\ \frac{1}{2} & \text{otherwise,} \end{cases} \quad (4)$$

It should be noticed that the *Tullock* form is discontinuous at the point of absence of fighting effort ($G_i = 0$, $\forall i$), in this case the usual splitting rule is that each Team wins with probability 1/2.

A second canonical CSF is known as the *logit* form and was originally suggested by Hirshleifer (1988, 1989). Here, the probability of winning of each team depends on the absolute difference in fighting efforts committed to the struggle :

$$p_1 = \psi_1(x_d) = \frac{1}{1 + \exp(m(G_2 - G_1))} \quad (5)$$

A third specification, the *probit* form from Baik (1998), also considers the absolute difference in fighting efforts as the main driver of the probability of success. This form holds that Team 1 wins the conflict if $mG_1 + \epsilon_1 > mG_2 + \epsilon_2$, where $\{\epsilon_i\}_{i=1,2}$ is a normally distributed noise component. Based on this, we get the *probit* form CSF as follows:

$$p_1 = \psi_2(x_r) = \Phi(m(G_1 - G_2)) \quad (6)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution as described in (2).

More recently, Beviá and Corchón (2015) proposed to consider the difference in fighting efforts relative to the total effort of both teams. By doing so, they capture the importance of the scale of the conflict. Their *RDCSF* defines Team1’s probability of winning such that:⁷

$$p_1 = \psi_3(x_{rd}) = \min \left\{ \max \left\{ \frac{1}{2} + m \frac{G_1 - G_2}{G_1 + G_2}, 0 \right\}, 1 \right\} \quad (7)$$

⁵ We are grateful to an anonymous referee for this suggestion.

⁶ For an illustration of the role of the decisiveness parameter in human warfare, see Hirshleifer (1995, pp. 44–46).

⁷ In their paper, Beviá and Corchón (2015) initially suggested the following form

$$p_1 = \begin{cases} \alpha + m \frac{G_1 - sG_2}{G_1 + G_2} & \text{if } G_1 + G_2 > 0, \\ \frac{1}{2} & \text{otherwise,} \end{cases}$$

where $\alpha \in [0, 1]$, m is the decisiveness parameter and s is a weight of the fighting effort devoted by Team 2. Beviá and Corchón (2015) show that the following constraint needs to be fulfilled: $2\alpha + m(1 - s) = 1$. For the purpose of comparison with other forms of CSFs, we assume $s = 1$ which involves $\alpha = 1/2$.

These four forms of CSFs have important theoretical specificities. For example, the *Tullock* form does not admit a symmetric Nash equilibrium without guns spending. The *Tullock* and the *RDCSF* forms are homogeneous of degree zero, the *logit* form is not.⁸ Therefore, choosing a CSF is not neutral and has a drastic impact on the prediction of models of conflict.⁹ One is well advised to empirically test different CSFs' predictions so as to choose the 'right' theoretical properties.

2.2. Empirical works on CSFs and known data issues

To the best of our knowledge, only two works compare existing CSFs in a warfare context, but their scope is limited to *Tullock* and *logit* forms. Firstly, Lotzin (2011) focuses on the Napoleonic Wars, using data from Smith (1998)'s book, compiling more than 2000 battles during the 1792–1815 period.¹⁰ The dependent variable is the French troops' victory. Fighting efforts are approximated by the size of the forces deployed in a given battle, for both *Tullock* and *logit* forms. Then, Lotzin uses the ratio of the correct predictions to the sum of all of them as a measurement of goodness-of-fit. He finds that a *logit* form is a better predictor of battle outcomes than the *Tullock* form. Secondly, Hwang (2012) offers a deeper investigation focusing mainly on two different periods: 17th-century European battles and World War II battles. For the first period, the dataset comes from Bodart (1908) and consists of 315 battles. The second dataset is compiled by the Historical Evaluation and Research Organization (HERO, 1983) and comprises 140 battles. Hwang (2012) uses a maximum-likelihood method to estimate CSF parameters and takes into account additional model selection criteria including the Akaike information criteria (AIC) and the Bayesian information criteria (BIC). According to his results, the *Tullock* form provides better predictions for 17th-century battles, but the *logit* form has a better goodness-of-fit for WWII battles. According to the author, the key explanation for this shift is the evolution of military technology: "The range of rocket missiles became longer and the advantage of electronics in the form of radar and improved radio communication was tremendous. Such improvements and refinements in weaponry, transport, and communications brought vast changes in the technology of conflict" (*ibid.*, p. 232).

These pioneering contributions are fundamental for the use of GVB models because they are the only empirical validation available in the literature. However, the use of historical battle-level data is also prone to several important criticisms. Notably, many difficulties emerge with respect to the estimation of CSFs' parameters. The first and major issue is related to the fact that "[e]stimating a contest model requires one to have measures of effort level of all contestants" (Jia et al., 2013, p. 219). But, how to measure fighting efforts on a (historical) battlefield? Newspaper articles or other contemporary sources from crisis regions discussing battles and casualties, usually, are not to be taken at face value, particularly when they are used to estimate the efforts and losses of either side of a conflict. Both sides have strong incentives to distort available data for propaganda purposes. Nations are built upon heroes, myths and legends which can be hard to disentangle from real historical facts. As a result, fighting efforts are mostly unobservable by historians and contemporary observers. In such conditions, the best estimator often is the number of troops involved. But this is a poor proxy for fighting efforts, for instance, because it neither precisely measures how well soldiers were armed, nor how well they were trained. Another issue related to the collection of data is the existence of an *ex-post* judgment bias: "the military historians of course knew the outcome of the battles when they made the codings" (Rotte and Schmidt, 2003, p. 180). As a result, they interpret evidence through their prior beliefs leading to measurement errors. Jia et al. (2013) also highlight a reverse causality issue in historical battlefield data. Indeed, battles are not independent one to another: a previous big victory (resp. defeat) may impact the troops' morale, and therefore the outcome of the next battle. In such conditions, it appears hard to establish a clear relation between fighting effort and probability of winning. But any attempt to empirically estimate CSFs has to address these difficulties.

Jia et al. (2013, p. 220) hold that "it would require an *unrealistic amount* of data to achieve any reasonable power in testing between them [i.e. different forms of CSFs] statistically" (emphasis is ours). Given the issues just outlined, one may add that it also takes an *unusually precise* dataset. That is exactly why we promote the use of data coming from virtual worlds.

3. EVE Online

This paper examines data from the virtual world of *EVE Online*. *EVE* was published by CCP Games in May 2003. It is a science-fiction themed virtual world. The user's main task is to compete with others in both economic and military ways. Whatever goals you set yourself, you will have to earn virtual money as a means for achieving them. You can do so either by violently appropriating what other users have, or by productive means.

When you log in, you freely navigate a space ship through a vast, three-dimensional universe. Unlike in other games, you do not build a small-scale economy or command armies from the perspective of an all-powerful economic planner or

⁸ Note that the analysis of the axiomatic and stochastic foundations of the CSFs transcends the scope of this article. On this topic, see Jia and Skaperdas (2012), Jia et al. (2013) or Cubel and Sanchez-Pages (2016).

⁹ We do subscribe to the idea of formal symmetry between explanation and prediction in economic models (Hausman, 2008; Hempel and Oppenheim, 1948). The CSF with the highest explanatory power therefore should also be the one with the highest predictive power. CSFs certainly have predictive power in *EVE*, as every *EVE* fleet commander preparing for battle will send scouts to assess the size of the opponent's army – in order to choose battles wisely or make last minute adjustments (get additional players to go online, change his party's fighting efforts by urging team members to choose more expensive yet more powerful ships, etc.) (*EVEUniversity*, 2018).

¹⁰ Lotzin (2011) estimates that the presence of Napoleon has a significant (positive) impact on the battlefield.

Fig. 1. Relative positioning of hisec, lowsec, and nullsec – featuring population averages and share of total kills in January 2011 (Guðmundsson, 2011).

general. Instead, you assume a first-person perspective and only control one space ship at a time - with thousands of other users simultaneously doing the same thing. Just as you “are” one particular person (your “character”) in other massively multiplayer games (e.g. *Second Life*), in EVE you basically “are” this space ship, whose actions you control via keyboard and mouse. In January 2011, i.e. the period for which we have data, EVE had around 400,000 active users and an average of around 30,000 concurrent users logged in at any time of the day. All of these users interact on the same server, so that every user is able to interact with every other user.

3.1. EVE's economics and politics

EVE features an almost completely user-run economy. It possesses its own currency, ISK (“InterStellarKredit”). The exchange rate between ISK and the euro in January 2011 was about 1 EUR = 19.5 m ISK (Mildenerberger, 2013b). More than 1.5 million individual trades are effected on EVE's markets on a daily basis. EVE's money supply M1 of 445 trillion ISK (23m EUR) may serve as a benchmark for the size of the economy. EVE's economy is stable, with only mild inflation and deflation (Guðmundsson, 2009; 2011). Although users typically only handle small amounts of money (in real life terms) in virtual worlds, virtual currencies do provide analogous economic incentives to real world currency (Castronova, 2001; 2008; Fiedler and Haruvy, 2009). This means that the things lost by a player during a conflict have a ‘real’ value.

EVE's political environment is tripartite. *High security space* (hisec) resembles a night watchman state. A computer-controlled police actively enforces a basic set of property rights. In *low security space* (lowsec) laws are only passively enforced. In *null security space* (nullsec) full-blooded anarchy reigns. For example, if you destroy the ship of a fellow user in hisec, police forces punish this offense by destroying your ship in return. Additionally, your character's security standing will be lowered. If a character's “security standing” falls below a certain threshold, he is no longer allowed to enter hisec. If you commit the same offense in lowsec, the security standing decrease is the only punishment. In nullsec there are no state-enforced legal rules whatsoever. The effects of this governance system are manifest. Although more than half of the characters “live” in hisec, more than half of the around 320,000 “kills”, i.e. ship destructions, that happened in EVE in January 2011 took place in nullsec (see Fig. 1).

Users are free to move around between spaces at any time. No taxes are levied in either of them. Whereas purely trade-minded and less experienced players typically stay in the relative safety of hisec, more experienced players as well as those willing to become tycoons or to claim their own territory move to nullsec.

3.2. EVE's conflicts

EVE is an “open world” or “sandbox”, meaning you can, in principle, fly your space ship wherever you want and fight whenever you want and whomever you want. Yet, conflict takes on two main forms. First, there is large-scale warfare for

territory between formal groups of users called “corporations” or “alliances”. It takes place in nullsec, for nullsec is the only region wherein users can officially claim territory. Nullsec also features the most resource-rich solar systems. Nullsec is vast but possesses little infrastructure, so it takes the cooperation of a considerable number of users to control territory and to profit economically from it. A straightforward form of division of labor characterizes such user alliances. Whereas *industrialists* specialize in resource mining, manufacturing, and transport, *warriors* secure the borders and protect the economic operations.

The other main form of conflict is *piracy* (Mildenerger, 2015). Living as a pirate is a distinct lifestyle. It mainly involves roaming around attacking other users' ships in order to destroy and pillage them. Pirates live off whatever their victims produce and transport. It is mostly individuals or small groups who execute acts of piracy. Regions exist where piracy is particularly predominant, namely along the trade routes leading from nullsec to the trade-hubs in hisec. Piracy is the reason why lowsec is conflict-ridden.

In the following, we will be exclusively focusing on battles that took place in nullsec, and which did not involve any non-player-controlled (NPC) characters (i.e. virtual space ships controlled by the game's software). These battles for territory and power in EVE's most anarchic region and between its militarily most well-organized groups are the closest analog to real world human warfare. For we look at two opposing groups of real people who fight each other at one particular location, with each person only controlling his own actions, and with the overall fighting effort being coordinated by military officers - with the sole difference that the battlefield and the weapons are virtual. Imagine *Braveheart*, but in the distant future. We also exclude practice fights within corporations or alliances, i.e. battles in which members of the same corporation or alliance fight each other exclusively.

The developers and users of EVE cherish the complexity and realism of fighting in EVE. Each user may choose his personal ship from more than a dozen ship classes, which range from small, fast, fragile, and cheap frigates to massive, slow, expensive, but robust capital ships which can be equipped with extremely powerful weaponry. Each ship class has its particular strengths and weaknesses. While, say, an individual battleship is typically superior to a frigate, groups of frigates can often take on more powerful opponents. There are even more classes of weapons in EVE than ships, ranging from electronic warfare, over lasers to missiles - again with particular strengths and weaknesses. Similarly an array of defensive systems is available (platings, counter electronic warfare modules, etc.). A look at the official EVE Forums on “Warfare and Tactics” (EVE, 2016) conveys a good impression of the art and science of choosing and equipping your ship for battle in EVE.

What ships, weapons, and defensive systems can be used by the individual character depends on that character's *skill points*. Users choose what exactly their character should be able to do among the options presented to them by EVE's skill system. Over time, characters acquire more skill points and thus become more and more apt at handling the more advanced economic and military equipment available. It takes about two years to train a comprehensively skilled fighting character (Guðmundsson, 2010b). This unusually long time (in comparison to other games) strengthens the validity of our data. As in real life, probability of success depends on how well trained in fighting a combatant is, with training taking time. Also, investing in military training comes with the opportunity cost of being a less skilled trader or producer.

4. Methodology

4.1. Advantages of virtual-worlds research

Virtual-worlds research combines the control of laboratory settings and the ecological validity of fieldwork (Fiedler and Haruvy, 2009). The rich data available allows us to zoom in on the phenomenon of interest, controlling for spurious factors, thus raising construct validity to levels usually seen only in experiments. For example, EVE's servers hold data concerning all the fights a user engaged in, the amount of material resources he invested, the size of his losses, the extent of damage he dealt to the opposing force, and so on. Consequently, we are able to measure the fighting effort devoted in a battle by each participant in great detail.

Ecological validity is high, as users do not fulfill given tasks but act in a natural habitat. This allows us to circumvent, for instance, effects typically seen in lab experiments like increased cooperation-levels in repeated experiments (Holmes, 1967). A major advantage virtual worlds have over laboratory experiments instituting conflict settings in particular (e.g. Duffy and Kim, 2005; Powell and Wilson, 2008; Smith et al., 2012), is that experiments often are not complex enough to produce a situation resembling a true conflict situation (Abbink, 2012). In comparison to researching conflict in the lab, virtual-worlds research is preferable ethically. Neither the researcher nor the subjects endanger themselves. Furthermore, the researcher does not have to engage in the morally questionable practice of incentivizing subjects to fight each other.

Analyzing conflict on the basis of the logged server data of a virtual world is able to address the notorious lack of objective evidence that is due to the chaotic and obscure nature of real-world conflict situations. For example, propaganda and manipulated statistics do not come into play in EVE. The problem with historians knowing outcomes of the battles when they compile their findings (Rotte and Schmidt, 2003) is eliminated in virtual worlds, as the battle outcomes are not publicly known. Also, virtual-worlds research features a high degree of inter-observer consistency. As the software gathers the observational data automatically, every researcher is confronted with the same data. As subjects in virtual worlds are in their natural habitats and unaware of being observed, they are not prone to testing biases of any kind. For the same

reasons, reactivity effects play no role. Finally, observer effects are excluded. As the researcher uses already existing data, his personal characteristics cannot influence its reliability.

EVE's data allows us to deal with issues relative to dynamic and feedback effects. A dynamic effect appears when battles are not independent as the outcome of previous battles has an influence on the fighting efforts devoted in the present one. It causes severe estimation concerns. In EVE, this effect is supposedly lower because battles are not traumatic.¹¹ As for the feedback effect, the worry is that once the outcome of a battle is pretty much clear, both sides might cut back their efforts substantially. In this sense the development of the battle might have a feedback effect on the effort exercised, which creates endogeneity issues. But contrary to many sport contests, for instance, reducing efforts once the winner is almost determined is not how EVE battles work. Firstly, if one side substantially reduces its effort, the possibility of being defeated by the weaker party remains. Secondly, in EVE, it is not only important who wins a battle, but also by how much. As in real life human warfare, what each team wants is to weaken its opponent as best as it can in each battle. Each hostile ship destroyed has to be replaced by the opponent. So this kind of feedback effect is unlikely to arise.

Last, according to Jia et al. (2013, p. 219) a good data set for estimating and comparing CSFs “requires that the structure of information available in contests should closely resemble the structure and requirements of information in theoretical models, i.e. both prizes and the characteristics of participants should be known *ex ante* by every player involved.” In EVE, all the information about which team brings which ships as well as how much they are worth on the market and how powerful they are is known to all contestants.

Apart from these rather specific advantages of using virtual worlds as a tool in conflict research, there are other, more general aspects of virtual worlds which underlie their “great potential as sites for research in the social, behavioral, and economic sciences” (Bainbridge, 2007, p. 472). Castronova (2008) stresses how virtual worlds have the potential to be the Petri-dish of the social sciences. Whereas, for instance, biologists study cells and can easily grow and manipulate many different cell-cultures at once in Petri-dishes, social scientists, for practical and ethical reasons, cannot just grow a society and test how it reacts to a natural disaster or how its economy performs when a war wreaks havoc – unless social scientists turn to virtual worlds, as for example the spreading of an infectious disease in the virtual world of the online game “World of Warcraft” shows. In this world, avatars contracted a virtual disease that was intended to only infect high-level avatars in a certain, clearly delimited region. But somehow the disease spread beyond the intended boundaries and attacked low-level avatars as well – that soon died afterwards, not without acting as incubators themselves. Medical journals reacted to this event and tried to learn lessons for real life (e.g. Balicer, 2007; Lofgren and Fefferman, 2007). While this virtual plague was an unplanned event, virtual worlds give scientists the possibility to simulate the effects of drastic, society-wide events. In collaboration with developers, social scientists can plan a disease, a natural disaster, or a war. They could even intentionally simulate the end of the world – i.e. the final shutting down of the virtual world's servers – to see whether, on the last day, chaos would reign, or whether everyone would gather around campfires telling stories about former times. While difficult to set up, virtual worlds feature the general possibility to stage and research such events.

4.2. Limitations

The main limitation of virtual-worlds research lies in its external validity. There are two major issues. First, one might think that virtual worlds like EVE overstate the prevalence of conflict. Second, one might think that we are estimating EVE's “battle algorithm” rather than a real CSF.

Selection bias is a first concern as regards the prevalence of conflict. Even if EVE features a smooth age distribution from 12 to 75 years (average age: 31), 95.7% of the users are male.¹² Still, this selection bias seems acceptable in our context, considering the distributions of age and sex of armies on real-world battlefields. A second concern is that EVE is a low-stakes environment – and low stakes are known to reduce economic motivations (Slonim and Roth, 1998) which, in turn, might favor violent behavior. Still, conflict in EVE is not costless. Across all the battles analyzed in this article, 487 billion ISK (approximately 248,000 EUR) worth of virtual objects were destroyed. Importantly, stakes for violent and productive behavior are both low within the virtual world, which is crucial for GVB approaches to properly model behavior. More generally speaking, as we are concerned with fighting itself, namely with how the fighting efforts devoted by the warring parties influence the probability of winning a battle, and not the *emergence* or the *extent* of violence, many issues that typically affect the external validity of virtual-worlds research on conflict can be put aside.

By contrast, given that EVE is a designed, fictional environment, there is an issue with respect to whether what we are ultimately measuring is the CSF of the game, i.e. the game's algorithm for conflict, rather than a “real” CSF. With respect to this issue it is important to highlight the similarities and differences between EVE, other video games, and real-world battles.

¹¹ We provide two formal robustness tests to this effect in the online appendix. Compare also Section 6.1 below.

¹² Yee (2006) finds that across virtual worlds, the biggest share of users (around 50 per cent) is full-time employed, with the second largest group being full-time students (around 20 per cent).

In EVE, *unlike* in other video games, there is no developer-determined “battle algorithm” implemented in the code which automatically calculates the outcome of a battle given the forces present. EVE is not an arcade game like *Space Invaders*, in which the only choices in fighting the invaders is to move from left to right and press the fire button. Such a game forces a particular CSF on the user. EVE is not a real-time strategy game like *Command and Conquer* either, in which you first build an army, then press “attack”, and then basically wait for whether your forces are stronger than the computer-controlled forces (as determined by the game’s battle algorithm). EVE possesses no CSF of its own. There is no “autopilot” for fighting, and the human factor is decisive.

As, for example, in a World War II battle, in an EVE battle two groups of people face each other, with each combatant possessing a certain type of weapon, certain fighting skills, etc. Moreover, each individual combatant can only influence his own actions. In EVE, this means that you only fly one space ship, and that you decide – for this space ship – whether and when to use which offensive and defensive weapons by individual mouse clicks and keyboard commands.¹³ To be sure, the characteristics, effectiveness, and relative strengths of particular weapons under ideal conditions are given in EVE and coded in the software. That is to say that, in EVE, we can tell that laser gun *x* does *y* amount of damage if fired at range *z*, if the target does not move, if the target has no shields activated, and so on. But this clearly does not amount to a CSF which is pre-determined by the developers.

This is first, because there is no disanalogy to the real world here. Whereas in the real world, when facing an opponent at mid to long range, a gun is superior to a knife *ceteris paribus*, in EVE artillery cannons are superior to pulse lasers under such circumstances. Importantly, people can *choose* whether to use guns/artillery cannons or knives/pulse lasers in battle, with all kinds of economic considerations framing and influencing this choice. If there were an absence of specific information on how different weapons do relatively speaking to each other, the information criteria present in any CSF would be violated. To make rational choices, fighters need to know their options. Most importantly, they need to know how much “bang for the buck” they can expect.

This is second, because ideal conditions almost never hold, as all kinds of tactical decisions influence the actual outcome. For instance, in EVE how much damage you do with a particular weapon depends on the relative speeds and directions in which you and your opponent move. Even if ideal fighting parameters for ships and weapons are known, they do not just determine battle outcome.¹⁴

There are yet other ways in which the human factor is decisive for the outcomes of battles in EVE. Users’ past choices in the game will determine what ships and weapons they can afford and use. In fact, while improving his character, every user will have to solve his own guns versus butter problem, in that some choices promise greater wealth (which is ultimately needed for powerful ships and weapons), whereas other choices favor an immediate increase in firepower.¹⁵ Also, within a battle, the efforts of the virtual generals, i.e. the “fleet commanders” are decisive. EVE allows for flanking, surprise attacks, guerrilla warfare as well as raiding tactics – and in large battles there are users whose sole job is to coordinate the efforts of all the users actually fighting, to manage the logistics support, and so on. Good choices will be entirely situational depending on factors beyond the users control but not determined by the game – like the number of the opponents forces, their specializations as regards ships and weapons, the terrain, and so on. The twelve-page long “Fleet Command Guide” (EVEUniversity, 2017a) allows a first glance at the complexity of EVE battles. Because EVE battles are so complex, there is competition among user alliances to secure the services of the best fleet commanders, as is seen on EVE’s recruiting forums (CCP Games, 2017).¹⁶

To conclude, as regards the relation of fighting efforts to battle outcome, it is closer to the truth to state that *nothing* but the environment’s physics and technology are determined by the developers. All of the aspects mentioned make EVE a particularly realistic game in which battle outcomes are unpredictable in the sense that human choice rather than a pre-determined CSF is decisive. On a more general note, consider that whenever we empirically measure a CSF based on some particular dataset, we are measuring “merely” the CSF of, say, basketball or of World War II. There is no context-free CSF. And concrete estimations always measure the CSF of something.

¹³ Some advanced users simultaneously but independently use two computers, with two distinct subscriptions to EVE to control two ships at once in an EVE battle.

¹⁴ Note that it is an explicit design policy by CCP Games to make EVE’s physics realistic, i.e. to do their best to make the virtual fights feel like fights in a space ship would feel in a distant future (Guðmundsson, 2010a, personal communication). For instance, given the vastness of space, it takes a long time in EVE to travel between different solar systems, or within solar systems, or to close the gap on an enemy as you need to do if you only have short range weapons. Because space is vast, EVE is a slow-paced game. Another example are realistic weapon physics. Whereas a missile fired in EVE has a certain travel time until it hits its target, laser weapons hit instantly. Thus, it might happen that the ship firing the missile is destroyed by its opponent carrying a laser weapon before the missile even reaches its target. Finally, there are no systematic distortions of the kind to be found in other games in that weaker characters are systematically favored (“boosted”) when fighting stronger characters, so as to make life easier for the former. Fighting parameters are the same for everybody.

¹⁵ EVE’s skill system, in which you can only train one skill at a time, but in which there are different skills for improving your efficiency in fighting or producing respectively underlies this trade-off (EVEUniversity, 2017b).

¹⁶ One thing which is beyond the user’s choice is war technology. Within the period examined, there was no technological progress in EVE. In general, new weapon and defense technologies in EVE are *developer-induced*. Thus, unlike in World War II in which the *warring parties’* efforts to introduce new technologies played a decisive role, with our data we cannot account for new technologies which change the form of the CSF or its parameters during the conflict. This is a weakness our approach shares with all those studies practicing a *battle-centric* analysis. However, in order to understand war in a static framework our approach is very robust. Not having to deal with constantly changing CSFs also seems analytically preferable if the main purpose is to compare different forms of CSFs. Note that while the technological ceiling is fixed by the developers, users constantly try to get closer to this ceiling.

5. Data and empirical strategy

5.1. Data

In order to estimate and compare the four forms of CSFs, we built an original database based on battles in EVE in January 2011. As the EVE server data we rely on contains no information about battles - like when a battle starts and ends, who is fighting whom together with whom, etc. - but exclusively lists data concerning individual kills (see Appendix A), we rely on an R script to identify independent battles in the pool of over 300,000 kills that took place in nullsec space, and which did not involve any non-player-controlled (NPC) characters or intra-group kills.

Starting with the first kill on January 1, at 00:00 a.m., the script individually tests for each kill whether other kills happened in the same solar system no more than 30 minutes later.¹⁷ If at least 2 kills happened within a given solar system within 30 minutes, we start out considering these kills to jointly constitute a battle.

For each battle thus identified, the characters participating in the kills are then sorted into two teams. If, for example, in one kill five characters jointly killed a victim, then the victim goes into one team (Team 1), and the five attackers into the other team (Team 2). For all the other kills in the same battle, the script then looks for other team members by matching who killed whom together with whom. Unless two clearly opposing teams with no joint members emerge, the battle is discarded. Also, all the battles in which one team only consists of one character, i.e. battles of the form 1 v X, are discarded. Overall, this battle identification strategy leads to 72.3% of the roughly 300,000 kills to be considered as battle kills in our estimations.

Once all the battles and teams are identified, the script calculates for each team member the value of all the ships he brought to the battlefield, and the value of all the ships lost. It does so chronologically, i.e. it allows for one player to bring a first ship to battle, to lose this ship, to come back with a second ship, and so on. These numbers are added up to give team values.

Our script initially identified 19,449 battles. However, we dropped 220 battles of our sample due to missing informations relative to important characteristics for our studies (missing values in ship investments, or in skills of players involved in a specific battle). Yet the data loss rate of 1.13% is low, and our final dataset still consists of 19,229 battles.

Our outcome variable, W , is a dummy equal to 1 if Team 1 wins the battle and 0 in case of defeat. As described above, EVE does not possess a developer-determined CSF of its own which is integrated in the game code and tells us who has won a battle. Therefore, in order to determine the winner of the battle, we compute for each team the relative loss ratio (hereafter 'RLR'), which is the value lost by each team (the value of all ships lost) in a battle divided by the total value invested by each team (the value of all ships brought to battle), as a proxy. The team with the lower RLR is considered victorious.¹⁸ We see in Table 1 that Team 1 wins 7.1% of the battles. This low number is due to the way we built our script. We first pick a kill at random, with the victim of this kill becoming the first member of Team 1. As the losing team typically will lose more ships than the winning team, it is more likely to choose a kill belonging to the losing team, which explains the low rate of success of Team 1.¹⁹

Our main variable of interest is the amount of fighting effort devoted by each team. Fighting efforts are directly observable in our dataset in the sense that we have the precise value (in ISK) of all ships brought to the battlefield by each team. To be consistent with previous notations, the total value of ships is labeled G_i , $i = 1, 2$. We see in Table 1 that G_2 is much bigger than G_1 . This is again related to how our script operates. It is reasonable to expect that Team 1 is more prone to be outnumbered.

The military power of an army does not only depend on the value of its weapons, but also on the abilities of the soldiers and on how experienced they are. Our dataset allows for refinements in that direction, as the fighting skills of each individual player involved in the battle are available. These fighting skills are scaled from 0 to 9, 0 being a very new player, and 9 a very experienced one. We thus compute the Average Skill Points (ASP_i) for each team. Our main variable capturing the difference in skill points of two contesting teams is a dummy variable, $DASP$, signaling the battle in which team 1 has higher average skill points than their opponent. In other terms, $DASP = 1$ if $ASP_1 > ASP_2$, 0 otherwise. By doing so we control for the effect that skilled users of expensive weaponry are more effective than amateurs. We also propose an alternative variable $RASP$ which is based on the ratio ASP_1/ASP_2 .²⁰ The fact that, on average, $DASP$ is inferior to 0.5 and $RASP$ is inferior to 1 is consistent with the fact that Team 2 is more likely to be the stronger team.

¹⁷ For technical reasons, battles in EVE do not span across multiple solar systems - or battles taking place in neighboring solar systems at the same time can be considered as independent battles in each solar system (CCP, 2014).

¹⁸ We choose this measurement instead of the absolute value lost on the battlefield in order to avoid misinterpretations of the outcome of a battle. Consider the following example. Team 1 brings two ships, each costing 3,000,000 ISK, and Team 2 brings three ships worth 10,000,000 ISK each. If Team 1 loses its two ships and Team 2 only one, it clearly appears that Team 2 is winning. However, if we consider the absolute value lost, we obtain that Team 2 is defeated (10,000,000 ISK > 6,000,000 ISK). With our measurement of victory we obtain 1 and 1/3 respectively for the RLR of Team 1 and Team 2. Thus, it provides a more appropriate picture of the outcome of the battle. Note that when CCP Games analyses the outcomes of battles in EVE, the developers rely on the same method as we do (CCP, 2014).

¹⁹ This phenomenon is particularly true considering that *unbalanced* fights represent 29.72% of the battles against 1.72% very *well-balanced*. As a result, the probability likelihood that the first kill picked belongs to the losing team is extremely high. See more on balanced versus unbalanced fights below.

²⁰ ASP_i can never be zero, as each character's skill points are strictly positive.

Table 1
Descriptive statistics.

	N	Mean	Std. Dev.	Min	Max
<i>W</i>	19 229	0.071	0.257	0	1
<i>G</i> ₁	19 229	3.64e+08	1.75e+09	14396.86	4.40e+10
<i>G</i> ₂	19 229	1.05e+09	6.26e+09	521973.9	2.10e+11
<i>DASP</i>	19 229	0.282	0.45	0	1
<i>RASP</i>	19 229	0.83	0.423	0.111	7

Table 2
Pure CSFs' estimation.

	Tullock	Logit	Probit	RDCSF
<i>m</i>	1.45*** (0.021)	2.13e–09*** (1.52e–15)	3.85e–10*** (7.19e–12)	0.5*** (1.39e–4)
Obs.	19 229	19 229	19 229	19 229
LL	–6 319.64	–9 258.03	–11 660.9	–6 612.73
AIC	12 641.28	18 516.06	23 321.8	13 227.45
BIC	12 649.14	18 516.06	23 321.8	13 235.32

Standard errors in parentheses with *, **, *** significant at 10%, 5%, and 1%, respectively.

5.2. Empirical strategy

In order to estimate the four alternative CSFs, let W_j be the success of Team 1 in the j th battle. W_j follows a Bernoulli of parameter p_{1j} and equals 1 if Team 1 wins battle j and 0 otherwise, with p_{1j} being the winning probability of Team 1 given by the *Tullock* (T), the *logit* (L), the *probit* (P), or the *RDCSF* forms such that:

Model (T): $p_{1j} = F[mx_r]$.

Model (L): $p_{1j} = F[mx_d]$.

Model (P): $p_{1j} = \Phi[mx_d]$.

Model (RDCSF): $p_{1j} = F\left[\ln \frac{\psi_3(x_{rd})}{1 - \psi_3(x_{rd})}\right]$, with $F[x] = 1/(1 + e^{-x})$ and where Φ is the standard normal cumulative distribution function.

In order to estimate the decisiveness parameter of these models, m , we perform maximum likelihood estimations. The models above deal with estimations of the “pure” CSFs, meaning that they strictly correspond to theoretical forms existing in the literature. It should be noted that – except when we are specifically looking at pure CSFs – we added an intercept and a variable accounting for the skills of players to offer a more fine-grained analysis.²¹

In order to compare the different forms of CSFs we proceed in two steps. First, we deal with nested CSFs, namely the *Tullock* and the *logit* forms. Providing comparison requires here a specific methodology based on the estimation of a more general form of CSF (see Appendix B). At the end of this first step, we are able to eliminate either the *Tullock* or the *logit* from the race. Second, we compare the three remaining CSFs using more traditional goodness of fit indicators: the log-likelihood (LL), the Akaike information criterion (AIC) and Bayesian information criterion (BIC).²² In particular, the model with the lowest AIC and BIC is considered the one which best fits the data. We also provide Vuong (1989) and Clarke (2003) selection model tests to offer more detailed results (see Appendix C). It should be noted that, to ease the reading, technical details about how we compare CSFs are only developed in the annexes mentioned above.

6. Results

6.1. Main results

We start by estimating “pure” CSFs as in (4), (5), (6) and (7). Table 2 shows that all forms of CSFs exhibit a positive and significant value for the decisiveness parameter m . This means that the more Team 1 is militarily superior in terms of value invested in the battlefield, the higher its probability to win the conflict. The interpretation behind this positive number is the following. On average a team always obtains a higher probability of success by increasing by one ISK the resources devoted to the battle. Another way to interpret this coefficient is to say that the “stronger” – in terms of fighting effort committed – always has a higher probability of winning the battle.

²¹ It corresponds to a slight modification of the models presented above. For example, for the *Tullock* form, we estimate $p_{1j} = F(\beta_0 + \beta_1 DASP + mx_r)$. The same logic is at stake for the other models.

²² We use the following definitions : $AIC = 2k - 2LL$ and $BIC = -2LL + k \ln(N)$, where k is the model degrees of freedom and N is the number of observations.

Table 3
Baseline results.

	Tullock	Logit	Probit	RDCSF
<i>m</i>	0.811*** (0.024)	6.10e–10*** (3.74e–11)	8.96e–11*** (7.97e–12)	0.455*** (0.006)
<i>DASP</i>	1.203*** (0.063)	1.505*** (0.059)	0.772*** (0.029)	1.187*** (0.063)
<i>Constant</i>	–2.473*** (0.048)	–3*** (0.047)	–1.737*** (0.020)	–2.452*** (0.048)
Obs.	19 229	19 229	19 229	19 229
LL	–3 778.35	–4 370.28	–4 486.07	–3 745.19
AIC	7 562.69	8 744.56	8 976.14	7 496.38
BIC	7 586.28	8 760.29	8 991.87	7 519.98

Standard errors in parentheses with *, **, *** significant at 10%, 5%, and 1%, respectively.

Regarding the comparison of the different forms of CSFs, one can easily see in Table A.3 that the *Tullock* performs better than the *logit* form ($\rho_a = 0.83$). Moreover, AIC and BIC indicators and specification tests (line (1) in Table A.4) conclude that the *Tullock* form offers the best fit. Pure estimations are interesting because they correspond to the strict definitions of the CSFs in the literature. However, they omit one important aspect: the experience of fighters.

In Table 3 we present our baseline results accounting for the average skill points of each team by integrating the dummy variable *DASP*.²³ Judging by the decrease of the AIC and BIC indicators, the addition of the *DASP* variable clearly improves the fit. It confirms the argument according to which the experience of players matters for the outcome of conflict. Unsurprisingly, the value of coefficients associated with *DASP* is systematically positive; meaning that, on average, being more skilled than one's opponent positively impacts the probability of success. One can also note the decrease of the decisiveness parameter (*m*): once we account for the advantage in terms of ASP, the difference in terms of value invested is less discriminatory. Regarding the choice of the more adapted form of CSF, we first notice that the *Tullock* is preferred to the *logit* form ($\rho_a = 0.829$). Then, Vuong and Clarke tests show that the *RDCSF* form provides the best predictions (see line (2) in Table A.4). This implies that the difference of fighting efforts does matter but needs to be controlled by the size of the battle. In other words, having a small advantage is more decisive for a small-scale conflict than for a large one.

In order to assess the robustness of our findings, we check for alternative specifications and samples. In all cases, we obtain very similar results. The decisiveness parameter is always positive and significant, a higher *ASP* for one team always turns into an advantage on the battlefield, and the *RDCSF* form appears to be the most adapted CSF. First, we present in Table A.6 estimations using the variable of relative *ASP* (*RASP*) instead of *DASP* to capture the difference in terms of average skill points in teams. Results are very stable.

Second, we use an alternative measurement for victory. Indeed, one shortcoming may come from the fact that having the lowest RLR may not be a sufficiently accurate indicator for victory. As an alternative, one might suggest, that the team who suffers more losses towards the end of a battle (e.g. because it is retreating and thus more disorganized) should be considered the losing team. Our alternative proxy is built on this suggestion. A team is considered as winning if it scores no less than 80% of the kills with respect to the last 25% of the kills in a battle. If no team reaches this threshold, the battle is considered a tie. In order to have a particularly clear picture of who won the battle, i.e. to have a particularly robust test, we decided to only include battles in which the last 25% of a battle's kills correspond to at least 10 kills. Put differently, we dropped all battles in which less than 40 kills occurred.²⁴ We obtain a sample composed of 294 battles, 96 of which are coded as ties and are thus eliminated. In the end, we have 198 remaining battles. Estimations provided in Table A.7 show results quite close to the baseline estimations.²⁵ Last, the value of the Clarke test statistics allows us to conclude that the *RDCSF* form still offers the best predictions (see line (4) in Table A.4).

In our third robustness check, we truncated our database by eliminating the first and the last 2.5% of kills. One can easily see in Table A.8 that results are quite stable regarding our baseline estimations.

Fourth, we checked for the importance of dynamic effects. We find that neither the average number of prior battles per team, nor the average historical win ratio per team affect the robustness of our results. Put differently, our baseline coeffi-

²³ We checked for multicollinearity issues regarding the link between our measurements of difference in fighting efforts x_k and *DASP*. First, correlation coefficients are very low. Second, removing *DASP* from baseline regressions involves a negligible impact both on the value of *m* and on standards errors associated.

²⁴ Otherwise, in small battles the last player killed would systematically determine the outcome of the battle, which is not realistic.

²⁵ As the RLR is based on a ratio measurement one might think that it artificially favors the *Tullock* form and the *RDCSF*. In the online appendix, we therefore provide another alternative measurement for victory: total annihilation. Team 1 wins if all ships of Team 2 have been destroyed. (It should be noted that this way to code a victory excludes the largest battles from the sample, as these are more likely to feature at least one survivor per team.) Our results indicate that the *Tullock* and the *RDCSF* still outperform the *logit* and the *probit* form. The Vuong test advocates the *RDCSF*, but the Clarke test supports the *Tullock* form. Still, we can consider that this mixed result confirms the superiority of the *RDCSF* for the full sample. This is because considering total annihilation leads us to consider a sample with smaller battles (average total ISK invested in the battlefield is reduced by –33.72%) and more homogeneous battles (the variance of total ISK invested is ten times smaller).

Table 4
Very well-balanced fights.

	Tullock	Logit	Probit	RDCSF
<i>m</i>	0.575 (1.512)	−4.17e−10 (3.14e−10)	−2.39e−10 (1.85e−10)	0.288 (0.759)
<i>DASP</i>	0.827*** (0.292)	0.847*** (0.168)	0.484*** (0.293)	0.827*** (0.292)
<i>Constant</i>	−1.775*** (0.189)	−1.803*** (0.19)	−1.072*** (0.103)	−1.775*** (0.189)
Obs.	330	330	330	330
LL	−153.98	−152.91	−152.97	−153.98
AIC	313.95	309.81	309.94	313.95
BIC	325.35	317.41	317.54	325.35

Standard errors in parentheses with *, **, *** significant at 10%, 5%, and 1%, respectively.

cients are not sensitive to taking into account players who are either traumatized or particularly motivated (“bloodthirsty”) because of (the results of) previous battles. Also, integrating dynamic effects changes nothing about the *RDCSF* outperforming the other forms.²⁶

Last, we changed the definition of a battle in our R script. Instead of starting to consider specific kills to be part of a new battle after 30 minutes have passed without any kill, we shortened the interval to 20 minutes and to 10 minutes. The rationale behind this is that the shorter the interval, the more distinguishable parts of a bigger battle (e.g. smaller skirmishes that accompany the warring parties getting into their chosen strategic positions) will count as individual battles, rather than as parts of the bigger battle.²⁷ Table A.9 specifically presents the results for the 20 min interval which are consistent with our other findings.²⁸

6.2. Specific results for (un)balanced battles

The fact that the *RDCSF* outperforms the other forms of CSFs emphasizes the importance of taking into account the scale of the conflict. In other words the link between fighting efforts and probability of success is not just a matter of differences (relative or absolute) but also depends on the size of the battle. This encourages asking if the degree of (un)balancedness also maps the technology of conflict. In order to explore this question more deeply, we explore two subsamples.

First, in Table 4 we only keep the very well-balanced fights, i.e. battles in which the stronger team did not invest 20% more than its enemy, and in which there is a less than 10% difference in the number of ships involved. It appears that such situations are quite rare in EVE and represent less than 2% of total battles. For very well-balanced fights, the value of the decisiveness parameter is not significant anymore. In other words, a difference in terms of ISK invested seems not to be crucial in the determination of the winner. We interpret this as a logical consequence of the selection of the subsample. By keeping only similar teams in terms of fighting efforts, the difference does not hinge on investment but rather on the ASP. In other words, in a battle composed by similarly strong contestants, the more skilled team is more likely to win. Regarding the comparison of CSFs, we see in line (7) of Table A.4 that specification tests are often non conclusive. For very-well balanced fights different CSFs are quite equivalent.

In Table 5, we consider fights in which the stronger team spends at least ten times more ISK than its enemy and teams bring a similar number of ships (plus or minus 10%).²⁹ These situations may be considered *unbalanced* fights, and are rather common in EVE (more than 25% of the total battles). We obtain similar results to the baseline estimations: *m* is positive and significant, *DASP* is positively related with the probability of success, and the *RDCSF* form performs better than other CSFs (see Table A.4).

6.3. Relative size matters

Our analysis compares the four main forms of CSFs existing in the literature. However, claiming that relative difference matters in military success also requires to prove that x_{rd} is preferred to both x_r and x_d in each family of CSFs.³⁰ Table 6 presents all possible combinations explored in this article, also including the Che and Gale (2000) CSF (hereafter ‘C&G’) and

²⁶ See the online appendix for the construction of the respective indicators and exact results.

²⁷ As is to be expected, the number of battles increases with a shorter interval. This effect is counterbalanced, however, by the fact that we discard all battles with less than two kills and with 1 v X opponents – which become more likely with a shorter time interval chosen.

²⁸ The results for the 10 min interval are also consistent with our overall findings. So are the results for the 15 min interval – with the small exception that for the 15 minutes interval, the Clarke test (but only the Clarke test) suggests a better fit for the *ratio* form.

²⁹ We focus on battles with a similar number of ships because a situation in which two very expensive ships are fighting 20 very cheap ships may not necessarily be considered an “unbalanced fight”, but might be a conscious strategic choice. In EVE, a group of cheap frigates and one expensive battleship often are on par in a fight. In order to avoid this issue, we select battles in which the number of ships the two teams bring is similar.

³⁰ Thanks to an anonymous referee for raising this important point.

Table 5
Unbalanced fights.

	Tullock	Logit	Probit	RDCSF
<i>m</i>	0.675*** (0.042)	1.67e–09*** (2.25e–10)	6.80e–10*** (6.23e–11)	0.451*** (0.011)
<i>DASP</i>	1.524*** (0.271)	1.951*** (0.254)	0.805*** (0.108)	1.493*** (0.276)
<i>Constant</i>	–3.305*** (0.187)	–4.139*** (0.196)	–2.148*** (0.07)	–3.187*** (0.185)
Observations	5 714	5 714	5 714	5 714
LL	–257.86	–303.69	–298.11	–253.1
AIC	521.72	611.37	600.21	512.2
BIC	541.67	624.67	613.51	532.15

Standard errors in parentheses with *, **, *** significant at 10%, 5%, and 1%, respectively.

Table 6
Combinations of CSFs.

	x_r	x_d	x_{rd}
$F[x]$	T	L	Model 1
$\Phi[x]$	Model 3	P	Model 2
$F\left[\ln\left(\frac{\psi_3(x_k)}{1-\psi_3(x_k)}\right)\right]$	Model 4	C&G	RDCSF

Table 7
Che and Gale (2000) CSF and Model 1–4 (baseline regressions).

	C&G	Model 1	Model 2	Model 3	Model 4
<i>m</i>	3.03e–10*** (3.58e–12)	2.395*** (0.063)	1.274*** (0.033)	0.426*** (0.012)	0.107*** (0.001)
<i>DASP</i>	1.203*** (0.063)	1.177*** (0.063)	0.581*** (0.032)	0.595*** (0.032)	1.296*** (0.061)
<i>Constant</i>	–2.473*** (0.048)	–2.439*** (0.048)	–1.329*** (0.024)	–1.363*** (0.023)	–2.661*** (0.047)
Observations	19 229	19 229	19 229	19 229	19 229
LL	–3 778.35	–3 740.94	–3 719.26	–3 761.56	–4 156.29
AIC	7 562.69	7 487.88	7 444.52	7 529.11	8 318.59
BIC	7 586.28	7 511.47	7 486.12	7 552.71	8 542.18

Standard errors in parentheses, with *, **, *** respectively significant at 10%, 5% and 1%.

Table 8
Comparison of nested forms (Baseline).

	Tullock v Logit	Logit v Model 1	Tullock v Model 1	Model 3 v Probit	Probit v Model 2	Model 3 v Model 2	Model 4 v C&G	C&G v RDCSF	Model 4 v RDCSF
ρ_a	0.829			0.843			0.614		
ρ_b		0.108			0.094			0.079	
ρ_c			–1.887			0.166			–0.024
<i>m</i>	0.032	0.274	7.697	0.029	0.195	1.18	6.33e–07	0.095	0.533

four models that have not been investigated at all in the literature yet. Baseline regressions for these five new CSFs are reported in Table 7.

To demonstrate that relative difference is the main driver in military success, we need to compare nested CSFs using the method described in Appendix B. Table 8 presents the results for baseline estimations. First, we compare relative difference and absolute difference. In particular, x_{rd} offers a better fit than x_d if and only if ρ_b is close to 0. One can easily see that this holds for all families of CSFs. Second, the value of ρ_c allows to decide between x_r and x_{rd} measurements. In particular $\rho_c = 1$ advocates for the ratio of fighting efforts, and $\rho_c = 0$ for their relative difference. We find that, for the three families of CSFs tested here, x_{rd} is preferred to x_r .³¹

A last question has to be addressed to provide a complete view: among forms of CSF that consider the relative difference in fighting efforts, does RDCSF offer the best fit? Table A.5 (columns 1 and 2) indicates that it is not the case. The Vuong and Clarke tests favor Model 1 and 2 relative to the RDCSF. This means that, if we explicitly do take into account relative

³¹ It should be noted that when we compare the Tullock CSF and Model 1, we obtain $\rho_c = -1.887$, meaning that the model advocated is clearly closer to Model 1. However, this result also suggests that considering $x_{rd} = (g_1^3 - g_2^3)/(g_1 + g_2)^3$ would be even better suited to our data.

difference to predict military success, Model 1 and above all Model 2 outperform the *RDCSF*. To the best of our knowledge, these forms of CSF have not been treated in the literature. Yet overall these findings empirically corroborate the view of Beviá and Corchón (2015) that “there is an advantage of being stronger; however, this advantage must be weighted by a measure of how large the conflict is”. They show that relative size matters for military success.

7. Conclusion

Our analysis shows that the *RDCSF* as suggested by Beviá and Corchón (2015) outperforms the three main competing CSFs as studied in the literature – i.e. the *Tullock* form, the *logit* form and the *probit* form. Our findings are very robust to changes in the proxy for victory, to the integration of the fighting skills of the opposing teams' members as further variable, and when only considering well-balanced or unbalanced fights. As economic predictions drastically hinge on the form of CSF used, as the *RDCSF* form has not been empirically compared to other CSFs at all, and as the literature largely relies on the *Tullock* form (Skaperdas, 1996; Tullock, 1980), this finding is of considerable impact. Claiming that relative size matters means that – in the case of low (resp. high) decisiveness of the technology of conflict – relatively more (resp. less) resources need to be devoted to guns to obtain a high probability of success. In particular, this implies that winning a battle in asymmetric warfare (characterized by a low *m*) requires a massive level of fighting effort. This is prone to cause more diversion of resources from productive activities (butter).

In general, our article encourages giving more importance to the role of the relative difference in fighting efforts to model the probability of success. In particular, our analysis highlights that $p_1 = \Phi[m(G_1 - G_2)/(G_1 + G_2)]$, where Φ is the standard normal cumulative distribution function, is a CSF that should be investigated more in the literature linking CSFs and armed conflicts. Outside the sphere of armed conflict, we think that the relative difference measurement in efforts should be more extensively used in other conflict situations like for example sport contests or innovation races.

As important as these theoretical points are, we believe that the methodological contribution of this paper is equally important. In a field of research which mostly lacks empirical testing, we are only able to offer as extensive and detailed a comparison as we rely on virtual-worlds data. Especially with respect to GVB models and conflict research – i.e. in an area in which reliable and sufficiently fine-grained data in large amounts is particularly hard to come by (given inherently chaotic real-world battlefields and the practical and ethical unfeasibility of lab experiments) – virtual worlds offer a great opportunity to empirically validate theories. This article illustrates how this promising approach can be put into practice.

Appendix A. General description of data

This paper essentially draws on data from three .csv-files made available by *CCP Games*. The first contains information on all the ship destructions that happened in January 2011 in EVE. An overview of the relevant data for this paper is provided in Table A.1. The second file contains information on important attributes for all the characters in EVE as described in Table A.2. A third file details the average market values of all virtual objects (including ships) in EVE in the examined period. Taking these three files together, we can develop an encompassing picture of the battles taking place in EVE and their economic effects.

Table A.1

Relevant contents of the file *kills.csv*.

Attribute	Typical value	Description
no. of kill	42	Each kill is attributed a consecutive number
place	30,001,409	The ID number of the solar system where the kill occurred
time	03.01.11 10:44	The time of the kill
victim	90,225,239	The ID of the character whose ship was destroyed
victim's ship	670	The ID of the ship that was destroyed
victim's alliance	99,000,198	The ID of the alliance to which the victim belonged
destroyer	756,884,476	The ID of the character who fired the last shot destroying the victim's ship
destroyer's alliance	100,958,673	The ID of the destroyer's alliance
destroyer's ship	346	The ID of the ship the destroyer flew
For each of the potential additional attackers besides the destroyer		that also participated in the kill
attacker's security status	−2.0	The security status of the attacker
attacker's alliance	1,000,958,673	The alliance ID of the attacker
attacker	472,395,793	The character ID of the attacker
attacker's ship	346	The ID of the ship the attacker flew
attacker's damage	123	The damage done by each attacker to the ship of the victim in absolute numbers
For each item that was in the ship of the victim (or fitted on it)		
item	266	The ID of the item
item's amount	400	The quantity of the item in question
item's amount dropped	200	The quantity of the item in question that was dropped after the kill, i.e. that was not destroyed due to the explosion of the ship

Table A.2Relevant contents of the file *characters.csv*.

Attribute	Typical value	Description
character	90,225,239	The character ID
skill points	02. [5 m to 10 m]	How many skill points the character currently has. Given in categories.

Appendix B. Procedure to compare nested CSFs

One cannot use a traditional goodness-of-fit indicator to compare nested forms (Jia, 2008; Jia and Skaperdas, 2012; Jia et al., 2013). For each family of CSFs $\psi_t(\cdot)$, $t = 1, 2, 3$, we need a special procedure to compare each measurement of difference in fighting efforts x_k , $k = r, d, rd$.

Ratio (x_r) v absolute difference (x_d)

We replicate here the solution proposed in Hwang (2009) and Jia (2010) consisting in estimating a more general measurement:

$$\frac{m}{1 - \rho_a} (G_1^{1-\rho_a} - G_2^{1-\rho_a}) \quad (8)$$

Note that for $\rho_a = 0$, (8) corresponds to x_d and for $\rho_a \rightarrow 1$, we obtain x_r . For example, in order to compare the *Tullock* and the *logit* forms, we estimate the following econometric model:

$$y_j \sim \text{Bernoulli}(p_{1j})$$

$$p_{1j} = F\left[\nu(G_{1j}^{1-\rho_a} - G_{2j}^{1-\rho_a})\right],$$

where

$$\nu = \frac{m}{1 - \rho_a} \quad \text{and} \quad F(x) = 1/(1 + e^{-x})$$

Results are presented in Table A.3. One can easily see that the *Tullock* form performs better than the *logit* form (ρ_a is closer to 1 than to 0) for almost all our regressions. The only exception is when we use an alternative measurement to code victory. This may be due to the little amount of data available for this case (we only have 198 observations here).

Absolute difference (x_d) v relative difference (x_{rd})

In this case, we estimate the following argument:

$$m \left(\frac{G_1 - G_2}{(G_1 + G_2)^{1-\rho_b}} \right) \quad (9)$$

Note that for $\rho_b = 0$, (9) corresponds to x_{rd} and for $\rho_b = 1$, we obtain x_d . For example, in order to compare the *Che and Gale* and the *RDCSF* form, we estimate the following econometric model:

$$p_{1j} = F \left[\ln \frac{\psi_3 \left(\frac{G_1 - G_2}{(G_1 + G_2)^{1-\rho_b}} \right)}{1 - \psi_3 \left(\frac{G_1 - G_2}{(G_1 + G_2)^{1-\rho_b}} \right)} \right]$$

Ratio (x_r) v relative difference (x_{rd})

Lastly, in order to compare the ratio and the relative difference in fighting efforts, we use a procedure very similar to Hwang (2009) by estimating

$$\frac{m}{1 - \rho_c} \left(\frac{G_1^{1-\rho_c} - G_2^{1-\rho_c}}{(G_1 + G_2)^{1-\rho_c}} \right) \quad (10)$$

Note that for $\rho_c = 0$, (10) corresponds to x_{rd} and for $\rho_c \rightarrow 1$, we obtain x_r .

We are now able to compare all the nested forms of CSFs.

Table A.3Comparison of the *Tullock* and the *logit* forms.

	(1) Pure	(2) Baseline	(3) Alternative ASP	(4) Alternative victory	(5) Truncated sample	(6) Alternative script	(7) Balanced	(8) Unbalanced
ρ_a	0.83	0.829	0.835	0.526	0.833	0.856	1.228	0.79
m	0.057	0.032	0.036	3.18e-05	0.035	0.042	52.6	0.014

Appendix C. Procedures to compare non-nested CSFs (Clarke, 2007)

In order to compare non-nested CSFs, we use two methods: the Vuong test (Vuong, 1989) and the Clarke test (Clarke, 2003). For both tests, the null hypothesis corresponds to a situation in which the two competing forms of CSF cannot be distinguished in terms of performance. If the Vuong statistic is positive and significant, it means that the first form of CSF tested fits better with the data than the second form. On the other hand, a negative and significant Vuong statistic involves that the second CSF performs better. The Clarke statistics has to be compared to the number of observations. If it is superior to half of the observations ($obs/2$) and is significant, it may be claimed that the first form of CSF fits better with data than the second one. These tests are consistent and unbiased but, as pointed out by Clarke (2007), the Clarke test has a greater power than the Vuong test when the distribution of individual log-likelihood ratios is leptokurtic. We show in the online appendix that in all cases, the Clarke test seems more appropriate regarding the distribution of individual log-likelihood ratios. Accordingly, values of Clarke tests appear in bold in Tables A.4 and A.5.

Table A.4
Vuong and Clarke statistics.

		Tullock v Probit	RDCSF v Tullock	Logit v Probit	RDCSF v Logit	RDCSF v Probit
(1)	VT	51.25*** (0.000)	–11.37*** (0.000)	30.91*** (0.000)	17.61*** (0.000)	59.62*** (0.000)
	CT	16 212*** (0.000)	2 594 (0.999)	16 638*** (0.000)	12 849*** (0.000)	15 932*** (0.000)
	Obs/2	9 614.5	9 614.5	9 614.5	9 614.5	9 614.5
	VT	13.42*** (0.000)	5.69*** (6.31e–09)	3.84*** (5.97e–5)	9.18*** (0.000)	14.17*** (0.000)
(2)	CT	13 787*** (0.000)	11 259*** (3.46e–125)	10 239*** (1.1e–19)	13 455*** (0.000)	14 062*** (0.000)
	Obs/2	9 614.5	9 614.5	9 614.5	9 614.5	9 614.5
	VT	13.55*** (0.000)	5.5*** (1.92e–08)	3.13*** (5.64e–4)	9.34*** (0.000)	14.31*** (0.000)
	CT	13 623*** (0.000)	11 531*** (1.65e–169)	9 464 (0.985)	13 468*** (0.000)	13 966*** (0.000)
(3)	Obs/2	9 614.5	9 614.5	9 614.5	9 614.5	9 614.5
	VT	0.16 (0.435)	0.07 (0.469)	0.67 (0.251)	0.04 (0.485)	0.17 (0.434)
	CT	113** (0.027)	118*** (0.004)	105 (0.217)	109* (0.088)	113** (0.027)
	Obs/2	99	99	99	99	99
(4)	VT	13.7*** (0.000)	5.61*** (1.04e–08)	30.2*** (6.85e–4)	9.26*** (0.000)	14.52*** (0.000)
	CT	12 395*** (0.000)	10 201*** (1.35e–123)	9 059*** (3.02e–10)	12 190*** (0.000)	12 668*** (0.000)
	Obs/2	8 652	8 652	8 652	8 652	8 652
	VT	11.01*** (0.000)	4.05*** (0.000)	3.41*** (0.000)	10.22*** (0.000)	12.28*** (0.000)
(5)	CT	12 927*** (0.000)	12 156*** (0.000)	9 384 (0.999)	13 348*** (0.000)	13 676*** (0.000)
	Obs/2	9 723.5	9 723.5	9 723.5	9 723.5	9 723.5
	VT	–0.518 (0.302)	–0.251 (0.599)	1.181 (0.119)	–0.546 (0.708)	–0.518 (0.698)
	CT	154 (0.897)	269*** (1.33e–32)	267*** (2.48e–31)	151 (0.945)	154 (0.897)
(6)	Obs/2	165	165	165	165	165
	VT	2.74*** (0.003)	1.727** (0.042)	–1.19 (0.117)	2.889*** (0.002)	2.925*** (0.002)
	CT	3 163*** (2.97e–16)	3 489*** (6.89e–63)	816 (0.999)	3 458*** (1.49e–57)	3 241*** (1.47e–24)
	Obs/2	2 857	2 857	2 857	2 857	2 857

(1) corresponds to the pure estimations, (2) is our baseline. Robustness checks are the next four: alternative measurement of ASP is (3), alternative measurement of victory is (4), the truncated database is (5) and the alternative script is (6). Last, estimations on very well-balanced fights and unbalanced ones are respectively labeled (7) and (8). p-values are in parentheses with *, **, *** respectively significant at 10%, 5% and 1%.

Table A.5
Comparison of CSFs using x_{rd} .

	RDCSF v Model 1	RDCSF v Model 2	Model 1 v Model 2
VT	−0.805 (0.789)	−4.202*** (1.32e−5)	−5.715*** (5.47e−9)
CT	8 296 (0.999)	5 495 (0.999)	6 673 (0.999)
Obs/2	9 614.5	9 614.5	9 614.5

Appendix D. Robustness checks**Table A.6**
Alternative measurement for ASP.

	Tullock	Logit	Probit	RDCSF
<i>m</i>	0.802*** (0.024)	5.97e−10*** (3.76e−11)	8.55e−11*** (7.98e−12)	0.454*** (0.006)
<i>RASP</i>	0.906*** (0.050)	1.149*** (0.047)	0.634*** (0.024)	0.888*** (0.05)
<i>Constant</i>	−2.382*** (0.045)	−2.9*** (0.043)	−1.715*** (0.019)	−2.36*** (0.044)
Obs.	19 229	19 229	19 229	19 229
LL	−3 802.69	−4 398.03	−4 494.11	−3 770.33
AIC	7 611.36	8 800.06	8 992.22	7 546.67
BIC	7 634.95	8 815.79	9 007.95	7 570.26

Standard errors in parentheses with *, **, *** significant at 10%, 5%, and 1%, respectively.

Table A.7
Alternative measurement for victory.

	Tullock	Logit	Probit	RDCSF
<i>m</i>	0.933*** (0.304)	1.53e−10*** (5.54e−11)	7.72e−11*** (2.67e−11)	0.479*** (0.090)
<i>DASP</i>	3.988*** (0.679)	4.254*** (0.674)	2.274*** (0.308)	3.981*** (0.681)
<i>Constant</i>	−3.39*** (0.598)	−3.662*** (0.592)	−1.932*** (0.246)	−3.381*** (0.599)
Obs.	198	198	198	198
LL	−48.98	−49.10	−49.47	−48.98
AIC	103.97	102.20	102.93	103.96
BIC	113.83	108.78	109.51	113.82

Standard errors in parentheses with *, **, *** significant at 10%, 5%, and 1%, respectively.

Table A.8
Truncating database.

	Tullock	Logit	Probit	RDCSF
<i>m</i>	0.805*** (0.025)	5.65e−10*** (3.85e−11)	8.08e−11*** (8.02e−12)	0.454*** (0.006)
<i>DASP</i>	1.169*** (0.066)	1.473*** (0.062)	0.757*** (0.03)	1.153*** (0.065)
<i>Constant</i>	−2.449*** (0.05)	−2.983*** (0.049)	−1.728*** (0.021)	−2.427*** (0.05)
Obs.	17 303	17 303	17 303	17 303
LL	−3 435.32	−3 3990.03	−4 084.79	−3 403.23
AIC	6 876.63	7 984.06	8 179.58	6 812.46
BIC	6 899.91	7 999.58	8 189.1	6 835.74

Standard errors in parentheses with *, **, *** significant at 10%, 5%, and 1%, respectively.

Table A.9
Alternative script (20 min).

	Tullock	Logit	Probit	RDCSF
<i>m</i>	0.636*** (0.021)	3.88e–10*** (3.04e–11)	7.72e–11*** (7.64e–12)	0.408*** (0.008)
<i>DASP</i>	1.847*** (0.059)	2.133*** (0.057)	1.099*** (0.028)	1.85*** (0.059)
<i>Constant</i>	–2.588*** (0.049)	–3.169*** (0.048)	–1.787*** (0.021)	–2.558*** (0.049)
<i>Obs.</i>	19 447	19 447	19 447	19 447
<i>LL</i>	–4 439.71	–4 891.94	–4 950.739	–4 394.07
<i>AIC</i>	8 885.48	9 787.88	9 905.48	8 794.14
<i>BIC</i>	8 909.11	9 803.63	9 921.21	8 817.76

Standard errors in parentheses with *, **, *** significant at 10%, 5%, and 1%, respectively.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.jebo.2018.08.012](https://doi.org/10.1016/j.jebo.2018.08.012)

References

- Abbink, K., 2012. Laboratory experiments on conflict. In: Garfinkel, M.R., Skaperdas, S. (Eds.), *The Oxford Handbook of the Economics of Peace and Conflict*. Oxford University Press, pp. 532–556.
- Aizenman, J., Glick, R., 2006. Military expenditure, threats, and growth. *J. Int. Trade Econ. Dev.* 15 (2), 129–155.
- Anderton, C.H., Carter, J.R., 2008. Vulnerable trade: the dark side of an edgeworth box. *J. Econ. Behav. Org.* 68 (2), 422–432.
- Anderton, C.H., Carter, J.R., 2009. *Principles of Conflict Economics. A Primer for Social Scientists*. Cambridge University Press.
- Baik, K.H., 1998. Difference-form contest success functions and effort levels in contests. *Eur. J. Polit. Econ.* 14 (4), 685–701.
- Bainbridge, W.S., 2007. The scientific research potential of virtual worlds. *Science* 317 (5837), 472–476.
- Balicer, R.D., 2007. Modeling infectious diseases dissemination through online role-playing games. *Epidemiology* 18 (2), 260–261.
- Beviá, C., Corchón, L.C., 2015. Relative difference contest success function. *Theory Decis.* 78 (3), 377–398.
- Bodart, G., 1908. *Militar-Historisches Kreigs-Lexikon V1: 1618–1905*. C.W. Stern Kessinger Publishing, Wien and Leipzig.
- Castronova, E., 2001. Virtual Worlds: A First-Hand Account of Market and Society on the Cyberian Frontier. SSRN Scholarly Paper ID 294828. Social Science Research Network, Rochester, NY.
- Castronova, E., 2008. A Test of the Law of Demand in a Virtual World: Exploring the Petri Dish Approach to Social Science. SSRN Scholarly Paper ID 1173642. Social Science Research Network, Rochester, NY.
- CCP, 2014. “The bloodbath of B-R5RB”, retrieved from <https://community.eveonline.com/news/dev-blogs/the-bloodbath-of-b-r5rb/>.
- Che, Y.-K., Gale, I., 2000. Difference-form contests and the robustness of all-pay auctions. *Games Econ. Behav.* 30 (1), 22–43.
- Clarke, K.A., 2003. Nonparametric model discrimination in international relations. *J. Conflict Resolut.* 47 (1), 72–93.
- Clarke, K.A., 2007. A simple distribution-free test for nonnested model selection. *Polit. Anal.* 15 (3), 347–363.
- Corchón, L.C., Serena, M., 2017. Contest theory: a survey. In: Corchón, L.C., Marini, M.A. (Eds.), *Handbook of Game Theory and Industrial Organization*. Edward Elgar Publishing.
- Cubel, M., Sanchez-Pages, S., 2016. An axiomatization of difference-form contest success functions. *J. Econ. Behav. Org.* 131, 92–105.
- Duffy, J., Kim, M., 2005. Anarchy in the laboratory (and the role of the state). *J. Econ. Behav. Org.* 56 (3), 297–329.
- Ehrenberg, R.G., Bognanno, M.L., 1990. The incentive effects of tournaments revisited: Evidence from the European PGA Tour. *Industr. Labor Relat. Rev.* 43 (3), 745–885.
- EVE, 2016. *Warfare & Tactics* - EVE Online Forums, retrieved from <https://forums.eveonline.com/default.aspx?g=topics&f=249>.
- EVEUniversity, 2017a. *Fleet command guide*. In *Uniwiki*, retrieved from https://wiki.eveuniversity.org/Fleet_Command_Guide.
- EVEUniversity, 2017b. *Skills and learning*. In *Uniwiki*, retrieved from <https://wiki.eveuniversity.org/Skill>.
- EVEUniversity, 2018. *Scouting*. In *Uniwiki*, retrieved from <https://wiki.eveuniversity.org/Scouting>.
- Fiedler, M., Haruvy, E., 2009. The lab versus the virtual lab and virtual field – an experimental investigation of trust games with communication. *J. Econ. Behav. Org.* 72 (2), 716–724.
- Frank, J., Öztürk, T., 2015. Conflict networks. *J. Public Econ.* 126, 104–113.
- Garfinkel, M.R., Skaperdas, S., 2007. Economics of conflict: an overview. In: Sandler, T., Hartley, K. (Eds.), *Handbook of Defense Economics*, 2. Elsevier Science B.V., pp. 649–709.
- Garfinkel, M.R., Skaperdas, S., Syropoulos, C., 2012. Trade in the shadow of power. In: Garfinkel, M.R., Skaperdas, S. (Eds.), *The Oxford Handbook of the Economics of Peace and Conflict*. Oxford University Press, pp. 585–610.
- Garfinkel, M.R., Skaperdas, S., Syropoulos, C., 2015. Trade and insecure resources. *J. Int. Econ.* 95 (1), 98–114.
- Guðmundsson, E., 2009. Quarterly economic newsletter – EVE online 2nd quarter 2009.
- Guðmundsson, E., 2010a. The economics of EVE. Interview. Reykjavik. August, 10, 2010.
- Guðmundsson, E., 2010b. Quarterly economic newsletter – EVE online 3th quarter 2010.
- Guðmundsson, E., 2011. Quarterly economic newsletter – EVE online 4th quarter 2010.
- Hausman, D.M., 2008. *Philosophy of economics*, Third ed. Cambridge University Press.
- Hempel, C.G., Oppenheim, P., 1948. Studies in the logic of explanation. *Philos. Sci.* 15 (2), 135–175.
- HERO, 1983. *Analysis of Factors That Have Influenced Outcomes of Battles and Wars: A Data Base of Battles and Engagements*. National Defense University Press.
- Hirshleifer, J., 1988. The analytics of continuing conflict. *Synthese* 76 (2), 201–233.
- Hirshleifer, J., 1989. Conflict and rent-seeking success functions: ratio vs. difference models of relative success. *Public Choice* 63 (2), 101–112.
- Hirshleifer, J., 1995. Anarchy and its breakdown. *J. Polit. Econ.* 103 (1), 26–52.
- Hirshleifer, J., 2000. The macrotechnology of conflict. *J. Conflict Resolut.* 44 (6), 773–792.
- Holmes, D.S., 1967. Amount of experience in experiments as a determinant of performance in later experiments. *J. Pers. Soc. Psychol.* 7 (4), 403.
- Hwang, S.-H., 2009. Contest Success Functions: Theory and Evidence. University of Massachusetts, Department of Economics Working Paper.
- Hwang, S.-H., 2012. Technology of military conflict, military spending, and war. *J. Public Econ.* 96 (1–2), 226–236.

- Jia, H., 2008. An Empirical Study of Contest Success Functions: Evidence from the NBA. Department of Economics, University of California Irvine Working paper.
- Jia, H., 2010. On a class of contest success functions. *B.E. J. Theor. Econ.* 10 (1).
- Jia, H., Skaperdas, S., 2012. Technologies of conflict. In: Garfinkel, M.R., Skaperdas, S. (Eds.), *Oxford Handbook of the Economics of Peace and Conflict*. Oxford University Press, pp. 449–472.
- Jia, H., Skaperdas, S., Vaidya, S., 2013. Contest functions: theoretical foundations and issues in estimation. *Int. J. Ind. Organiz.* 31 (3), 211–222.
- Konrad, K.A., 2009. *Strategy and Dynamics in Contests*. Oxford University Press. LSE Perspectives in Economics Analysis
- Lastowka, F.G., Hunter, D., 2004. The laws of the virtual worlds. *Calif. Law Rev.* 1–73.
- Lessig, L., 1999. *Code and Other Laws of Cyberspace*. New York: Basic Books.
- Lofgren, E.T., Fefferman, N.H., 2007. The untapped potential of virtual game worlds to shed light on real world epidemics. *Lancet Infect. Dis.* 7 (9), 625–629.
- Lotzin, F.C., 2011. *The Emperor on the Battlefield: Napoleon's Worth as a Military Commander*. Grin Verlag.
- Ludlow, P., 2001. New foundations: On the emergence of sovereign cyberstates and their governance structures. In: Ludlow, P. (Ed.), *Crypto Anarchy, Cyberstates, and Pirates Utopia*. The MIT Press, Cambridge, MA, pp. 1–23.
- Maloney, M.T., McCormick, R.E., 2000. The response of workers to wages in tournaments: evidence from foot races. *J. Sports Econ.* 1 (2), 99–123.
- Mildenerberger, C.D., 2013a. The constitutional political economy of virtual worlds. *Constit. Polit. Econ.* 24 (3), 239–264.
- Mildenerberger, C.D., 2013b. *Economics and Social Conflict. Evil Actions and Evil Social Institutions in Virtual Worlds*, Economics & Finance Collection. Palgrave Macmillan.
- Mildenerberger, C.D., 2015. Virtual world order: the economics and organizations of virtual pirates. *Public Choice* 164 (3–4), 401–421.
- Pareto, V., 1971. *Manual of Political Economy*. London Macmillan.
- Peeters, T., 2011. The Shape of Success: Estimating Contest Success Functions in Sports, Working Paper Series, Paper No. 11-08. International Association of Sports Economists.
- Powell, B., Wilson, B.J., 2008. An experimental investigation of hobbesian jungles. *J. Econ. Behav. Org.* 66 (3), 669–686.
- Robinson, N., 2016. Militarism and opposition in the living room: the case of military videogames. *Crit. Stud. Secur.* 1–21. doi:10.1080/21624887.2015.1130491.**
- Rotte, R., Schmidt, C., 2003. On the production of victory: empirical determinants of battlefield success in modern war. *Defence Peace Econ.* 14 (3), 175–192.
- Skaperdas, S., 1996. Contest success functions. *Econ. Theory* 7 (2), 283–290.
- Slonim, R., Roth, A.E., 1998. Learning in high stakes ultimatum games: an experiment in the slovak republic. *Econometrica* 569–596.
- Smith, A.C., Skarbek, D.B., Wilson, B.J., 2012. Anarchy, groups, and conflict: an experiment on the emergence of protective associations. *Soc. Choice Welfare* 38 (2), 325–353.
- Smith, D.G., 1998. *The Greenhill Napoleonic Wars data Book*. Greenhill Books, Mechanicsburg, PA: Stackpole Books, London.
- Sunde, U., 2009. Heterogeneity and performance in tournaments: a test for incentive effects using professional tennis data. *Appl. Econ.* 41 (25), 3199–3208.
- Tullock, G., 1980. Efficient rent seeking. In: Buchanan, J.M., Tollison, R.D., Tullock, G. (Eds.), *Towards a Theory of the Rent-Seeking Society*. Texas A&M University Press, College Station, pp. 97–112.
- Van Long, N., 2013. The theory of contests: a unified model and review of the literature. *Eur. J. Polit. Econ.* 32, 161–181.
- Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 307–333.
- Yee, N., 2006. The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence* 15 (3), 309–329.
- Yildizparlak, A., 2017. An application of contest success functions for draws on European Soccer. *J. Sports Econ.* (Forthcoming)