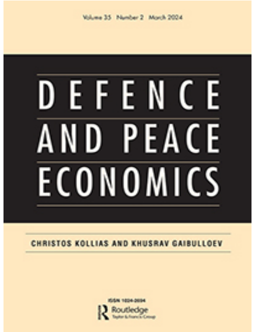

Defence and Peace Economics

ISSN: 1024-2694 (Print) 1476-8267 (Online) Journal homepage: www.tandfonline.com/journals/gdpe20

The Political Economy of Wargaming

Garrett R. Wood

DEFENCE AND PEACE ECONOMICS
2024, VOL. 35, NO. 2, 160–173
https://doi.org/10.1080/10242694.2022.2145537

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

The Political Economy of Wargaming

Garrett R. Wood

Department of Management, Business, and Economics, 5817 Wesleyan Drive, 23502, Norfolk, United States

ABSTRACT
Modeling conflict through wargaming is the only option outside of high-cost real conflict for militaries to train their forces and attempt to reveal information about their own strengths and weaknesses as well as those of their foes. This is the function wargaming serves in theory, but in reality, the process of wargaming is undermined by information and incentive problems that cause the real-world performance of wargames to deviate sharply from their performance in theory. These problems resolve the conflicting professional views on wargaming between those who want to use them for predictive purposes and those who want to use them for training purposes in favor of the latter.

ARTICLE HISTORY
Received 19 July 2022
Accepted 6 November 2022

KEYWORDS
Wargaming; national defense; information problems; incentive problems

JEL CLASSIFICATION
D83; F50; H56

Section 1. Introduction

In the span of ten minutes an overwhelming missile attack destroyed one US Navy aircraft carrier, five amphibious assault ships, and several cruisers, wiping out an entire invasion force with more US lives lost than in decades of war in Afghanistan. Senior U.S. military officers were stunned, but they soon reset the game and handicapped the team playing the enemy forces to prevent them from defeating the simulated U.S. forces a second time. U.S. success after the reset of the game was then used by the Department of Defense to validate certain operational concepts and encourage the purchase of weapons systems that would support them before Congress (Zenko 2015). Wargaming is one of the few options available to militaries to safely test ideas, leaders, and platforms outside of real war, so the failure to use them for these ends can result in lethal mistakes during real wars. This paper seeks to answer the question of whether wargames are appropriate for predictive purposes or are better used as intuition training devices by applying an economic analysis of information and incentive problems within the practice of wargaming.

Wargames must be judged by their robustness against conditions that are not ideal because these are the conditions found in real-world practice. To illustrate by comparison, prices continuously serve to signal underlying changes in market conditions while incentivizing the appropriate responses to those changes. This makes prices robust in their ability to promote the efficient use of scarce resources even when assumptions of perfect information and benevolence are relaxed. No similar mechanism exists for wargaming, meaning that when assumptions of perfect information and benevolence are relaxed it becomes far more likely that a wargame will play out similarly to the wargame that opened this paper. War itself ultimately serves as the relevant feedback mechanism, but its information revealing properties and incentive structures are not present during peacetime wargaming.

This economic insight into wargaming is a unique addition to the wargaming literature. The existing literature is split between focusing on attempting to make wargames produce reliable and

CONTACT Garrett R. Wood garrettrwood@gmail.com Department of Management, Business, and Economics, 5817 Wesleyan Drive, 23502, Norfolk, Virginia

reproducible quantitative data for predictive purposes (Hernandez, McDonald, and Ouellet 2015; Work and Selva 2015; Appleget and Cameron 2015; Sloan 2020), and focusing on using wargames as intuition training devices for military personnel (Beuch 1987; Perla 2008; Pournelle 2016; Barney 2020). This paper argues that the split between these two approaches must be settled in favor of the latter approach, not because this approach is free of the information and incentive problems that plague wargaming, but because it is comparatively less affected by them and less likely to be used to intentionally generate misleading results. It analyzes the practice of wargaming through the lens of political economy. Even when the incentives of wargaming decision-makers are aligned with the interests of society, they may suffer from problems of incomplete information that prevent them from making optimal decisions (Hayek 1945; Kirzner 1984; Lavoie 1985 [2016]). Problems of incomplete information in the arena of military affairs have been explored in a variety of settings (Olson 1962; Coyne 2015; Wood 2018; Coyne and Hall 2019). Additionally, decision-makers involved in ordering and running wargames make decisions based on the incentives they face within their political institutions, meaning their actual behavior may diverge from that which would maximize the public's consumption of national defense (Brennan and Buchanan 1985; Buchanan 2003; Acemoglu and Robinson 2013). Agents pursuing their own interests at the expense of the public's consumption of national defense has been covered in a variety of settings (Melman 1985; Brandes 1997; Punjani 2012). I use these two sources of government failures to explain the mechanisms and incentives behind wargaming failures.

I draw on work done by military scholars publishing at institutions of higher military education as well as private scholarly outlets, on declassified wargame reports, on Government Accountability Office reports, and on investigative journalism to illustrate the information and incentive problems in wargaming. The paper will proceed as follows. Section 2 will describe how information and incentive problems cause failures in wargaming. Section 3 will explore the incidence and persistence of these failures from real-world wargames. Section 4 will conclude.

Section 2. The Purpose of Wargaming and Its Limits

A wargame is an attempt to model the most relevant characteristics and behaviors of one's own forces, typically known as Blue forces, and those of one's enemy, typically known as Red forces. Participants are tasked with commanding simulated Blue forces and simulated Red forces during an engagement or engagements with each other. The agents running the game are known as White Cell and are generally mid-level to senior officers depending on the scale of the game. They have discretion over in-game resources, rules of engagement, rules governing unit behavior and interactions, and most other elements of the game. This discretion gives them a variety of margins on which to rig the results of the wargame, and their potential personal gains from doing so will be described later in this section. These agents report to a principal who ordered the game and sets broad parameters for what the game is intended to study. Depending on the scale of the game this principal would be a senior officer or senior civilian bureaucrat, or even a top-level politician such as the President of the United States who requires the information produced by the wargame to make decisions about matters such as military funding, deployments, tactics, and strategy. The more accurate the information the game produces, the better the decisions made by the principal. Since the reverse is also true, principals would prefer that their agents running the game do so honestly.

Wargames can be single shot, repeated, or can be chained together such that the outcome of a game informs the start point of a follow-on game. Wargames are capable of modeling multiple scales of conflict and differing levels of complexity at relatively low cost compared to war because they are typically run on a tabletop or simulated on a computer. However, the cost of incorporating more detail is still positive and increasing. These information costs combined with the incentives agents have to pursue their self-interest cause wargaming to not be robust in its ability to create useful predictions or discover relative strengths and weaknesses under real-world rather than theoretical conditions. The way these two problems manifest in wargaming will now be described in further detail.

The first problem wargaming faces is an information problem. This breaks down into problems of the quantity of information necessary for faithfully reflecting reality and of decentralized information held by actors other than the agents running wargames. These problems mirror the problems in economic calculation but are made more severe by the lack of prices, profit, and loss in non-market systems of resource allocation (Hayek 1945; Kirzner 1984; Lavoie 1985 [2016]). Prices reduce information about input scarcity, production costs, opportunity costs, and preferences down to a simple number. No similar mechanism exists in wargaming for reducing information about the value of military labor and capital in alternative uses, and the failure of attempts to find such a mechanism is illuminating. For instance, the wargame could attempt to measure success in terms of effects against critical resources, but this proved unreliable in World War II (Olson 1962), it could attempt to measure success in terms of attrition against enemy capabilities, but enemy capabilities dynamically respond to shifting constraints produced by combat (Wood 2018), or it could attempt to measure success in terms of body counts, but this proved very misleading during the Vietnam War (McMaster 1997). Without a mechanism for economizing on the quantity of information necessary for military planning and gaming, agents have to resort to direct consideration of the scarcity of every input and its possible alternative uses (Neurath 1919). The quantity of information problem causes accuracy in wargaming to be very costly to achieve and causes the communication of wargame results to be very costly as well. These high information costs undermine the effectiveness of wargames, and especially wargames as high-fidelity predictive models.

For wargames to be practical, those who run them must abstract away from complexity on a multitude of margins. These could include both difficult to quantify variables such as morale and easily quantifiable variables such as the reliability of radio communications, but some level of abstraction is necessary to mitigate the quantity of information problem. The need for abstraction and simplifying assumptions only increases when it is recognized that a game must be run multiple times, with limited resources, if the goal of the wargame is to produce reliable predictions (Anderson and Dilegge 2015). The cost in terms of time can quickly become too high for militaries already burdened with numerous other training requirements. Examples of honest mistakes caused by the need to abstract from the large quantity of information necessary for wargames run for predictive ends could include using outdated information about logistical infrastructure, using a unit's manning and equipment on paper without confirming either in person, using the wrong operating procedures for a modeled unit due to lack of familiarity, or other similar issues. Further complicating matters, in scenarios that incorporate civilian populations or infrastructure the military will have to acquire necessary information from civilians not within its command structure. This increases the costs of acquiring relevant information even when civilians want to cooperate and adds to the decentralized information problem. The information problem in wargames exists, and must always exist, in part because the information necessary to run the game is decentralized across a variety of individual actors both in and out of the agency responsible for the wargame.

The problem of decentralized information also extends to the information a wargamer needs to accurately model their opponents, which is held secret by those opponents. Indeed, private information is considered to be one of the main reasons for the failure of belligerents to reach a pre-war bargain which is preferable to the post-war division of spoils (Fearon 1995). Actual warfighting incentivizes belligerents to reliably reveal their private information, but this incentive is missing in wargaming. The agents running the game must rely on military intelligence and reasonable guesswork regarding the capabilities and disposition of the Red Forces they model. The scope for honest mistakes here is obvious, an agent could fail to incorporate accurate information on things such as the durability of upgraded enemy armor or the development of innovative new enemy tactics which have not yet been revealed. The potential for these mistakes is increasing with time between wars because technological innovation is continuous, and tactical or strategic innovation may also be continuous. Taken all together, the information problems wargaming faces are enough to significantly undermine its ability to be used to make accurate predictions, but that also create an environment in which the second problem wargames face can arise.

The second type of problem wargames face is incentive problems. The public sector agents who develop, run, and report the results of wargames are self-interested and may pursue those interests over the interests their principals have in generating high-quality information suitable for military planning (Brennan and Buchanan 1985; Buchanan 2003; Acemoglu and Robinson 2013). The information problems discussed above would be enough to severely compromise the validity of wargames as planning tools, but they also contribute to an environment with a high noise-to-signal ratio in which agents can hide their actions and true intentions. The agent's pursuit of their self-interest can be covered up using problems relating to the volume of information or decentralized information as excuses, leaving both expert and non-expert principals in an environment with high monitoring costs. The worse the information problems are, the more difficult detection becomes.

Agent interests diverge from their principal's interests for a variety of reasons. They may secure kickbacks from defense firms who would otherwise lose government contracts in exchange for rigging a wargame to paint that firm's product in a positive light relative to competing products, their careers may be tied to certain weapons systems that they can hide the flaws or obsolescence of by rigging wargames, their promotion or the promotion of favored subordinates may be tied to the results of a wargame which they would then have an incentive to rig, or they may even rig wargame results to avoid scandals and budget cuts resulting from ensuing political or civilian oversight. The margins on which agents can attempt to rig their games are numerous. They could artificially constrain enemy forces to certain activities, give friendly forces next-generation technologies while holding enemy forces to current generation technologies, inconsistently apply metrics of effectiveness, artificially constrain enemy force deployment areas, intentionally neglect to check assumptions about supporting infrastructure or logistics, or use a variety of other methods. The complexity of conflict and lack of a mechanism that economizes on information, such as the price mechanism, almost guarantees the existence of difficult-to-monitor margins on which self-interested agents can rig their wargames. While actual conflict is likely to reveal the margins on which deception occurred, sufficiently low-intensity conflict may significantly delay this revelation, or the agent could reasonably claim that the parameters of actual conflict were different from those he was tasked with modeling, or the complexity of war may cause the outcome to be incorrectly attributed to some other element co-producing defense with the rigging agent. This problem is mitigated when the wargame models an ongoing conflict with the accurate information produced by that conflict. In this case, rigging efforts will be more quickly revealed by the unfolding of the actual conflict. When wargames are run to test operational concepts and technologies in hypothetical wars, however, that real-world auditing function performed by war is absent and detection of rigging becomes less likely.

Both information and incentive problems grow with the scale of the wargame. Increased scale increases the quantity of data necessary to run the wargame and the likelihood that relevant information will be held by actors other than those ordering and running the wargame. Likewise, the larger the scale of the wargame the more weapons systems and communities are likely to be involved, necessitating a larger White Cell and therefore an increase in the number of self-interested agents. The effects of information and incentive problems in wargaming limit the usefulness of wargames to being intuition training devices, not unlike the way chess has been used to train the intuition of military leaders who will never meet a rook in actual battle, rather than being reliable predictive models. This is not to say that wargames never make accurate predictions, only that they cannot be relied on to do so and that reliance on wargames for this purpose enables self-interested agents to influence decisions about defense in their own favor at the expense of their principal's interests and the optimal provision of national defense.

The effects of information and incentive problems in wargaming could be dismissed as anecdotal, but their persistence suggests that they are not accidents of circumstance and are instead systemic problems. Mistakes could eventually be worked out by institutions with long time horizons such as the Department of Defense, but the repetition of errors and intentional rigging suggests that these institutions are not able to overcome the information and incentive problems inherent in wargaming.

Section 3. Information and Incentive Problems in Wargaming

Section 3.1. Information Problems

The information problem arises from the fact that much of the information necessary for wargaming is decentralized and enormous in quantity. Necessary information is held by a variety of civilians, a variety of communities within the military, and a variety of potential enemies. The communication of that decentralized information is costly, and those costs increase with the sheer volume of information necessary to run wargames. Information problems manifest in a variety of ways and at all levels of wargaming. This section will explore the decentralized information problem first and the difficulty of economizing on and communicating large quantities of information second.

Section 3.1.1. Problems of Decentralized Information Held by Civilians and Blue Forces

The scope and complexity of the information problem is lessened when considering only the information which is needed to model Blue Forces. This is because, in theory, all of that information is known to individuals friendly to Blue Forces who would share it. The wargame Nifty Nugget was run in 1978 to test the ability of the United States to mobilize and deploy forces to Europe in response to Soviet aggression. It was predominantly a test of US mobilization capacity and was therefore focused on information known to individuals within the United States. On paper, Blue Forces had sufficient logistics capability to not only move hundreds of thousands of troops to Europe but to supply them while evacuating a million US citizens from the combat zone. In playing the game, however, they discovered that simultaneous logistics support of so many troops while evacuating US citizens was impossible (Clem 1983). Worse, the information on civilian logistics capabilities Blue Forces relied on were often out of date. Rail lines and airports no longer existed, had been moved, or were sufficient only for transporting older versions of weapons that were smaller and lighter. This was particularly the case with the rail system that Blue Forces were supposed to use, with lines having long ago been shut down or fallen into disrepair or occupied by cars that could not physically move modern tanks. Information about altered rail services would have been communicated to commercial users through the price system, but military users are not fully integrated into this system. The result of these failures was that some 400,000 troops were killed or wounded within just a few weeks because of their lack of fresh food, ammunition, fuel, and spare parts shipments. An institution operating as a singular organism would have simply known where its logistics assets were, and what capacity they held, only missing critical information held by its opponents.

As a singular example, it could be argued that this is simply an instance of wargames performing their intended function of revealing information and overcoming the information problem. The wargame Proud Spirit was run in 1980 as a follow-up to Nifty Nugget and was meant to demonstrate that the lessons of Nifty Nugget had been learned. The scenario being tested was identical save that there was no conflict to fight or refugees to evacuate, only a mobilization of troops and resources to Europe. This means the information problems were simplified and should have resulted in success. Instead, military planners were again caught off guard by changes in rail location and capacity, as well as changes in airport capacity (Clem 1983). These problems would have reproduced the large-scale death and defeat seen in Nifty Nugget had the combat portion of the game not been removed, and they show that the decentralized information problem is a persistent one in wargaming.

The problem of decentralized information applies when dealing with a military that is partially reliant on outside civilian capabilities, but it also applies purely within the military. Not only are militaries typically broken up into individual branches such as those commanding ground or sea forces, but there are also further divisions representing different communities within those branches. For instance, the US Navy contains different communities for aviation, surface ships, subsurface ships, riverine forces, and more. Each branch, each community, and each further subdivision holds decentralized information about what assets they have, what condition they are in, and what human capital they possess. For this reason, the Strategy and Force Evaluation games run throughout the 1960s were impaired in their ability to make useful predictions because they were played by Air Force officers without input from Navy officers. The result of not having access to the information held by experts in naval warfare was that the Air Force officers playing the game repeatedly predicted massive build-ups in land-based mobile nuclear missile launchers by both the US and the Soviet Union. In reality, both militaries pursued nuclear-armed ballistic missile submarines instead, a prediction the game players would have likely made had they reached across the Department of Defense to include Navy officers with valuable decentralized knowledge about these comparatively long-range and elusive assets (Brown and Paxson 1975).

A further complication for militaries dealing with the decentralized information held by their own branches and subcommunities is the fact that aggregating that information does not necessarily create an accurate picture of combat. One of the foremost figures in American wargaming, Trevor Dupuy, lamented the inability of his rigorous and computerized wargaming models to deliver intelligible aggregated results. Allen (1987 pg. 254) quotes him as saying: 'We have not yet got to the point that we can relate our battalion model to our brigade models, to our division models, to our theater models. We can't take the results of the lower-level models and make them coincide with the results of the higher-level models. You run into very difficult technical and conceptual problems. At what level do you model the Vietnam War? If you model it at the highest level, you can't relate that to the results of a company or a battalion level engagement.' Another example of this problem in wargaming is discussed in Allen (1987 pg. 252) as he quotes Major General Edward Atkeson of the Army's Concepts Analysis Agency: 'You ambush the other guy. And then you'll get very favorable exchange ratios. Well, the Army wants to have consistency. We're teaching guys down at our platoon and company level – and what happens at battalion? You get a whole bunch of ambushers. Well, what happens at brigade – and corps? And Theater? You end up with a World War Three that is a thousand ambushes. It means that somehow we have to find a way to fight World War Three where we pick the terrain. We pick the phase of the moon, the time of the day, the season of the year. Well, that's not the way the enemy usually likes to play.' Even wargames that accurately reflect the decentralized information held by the various communities involved struggle to combine that information into a cohesive framework for testing higher-level engagements as the scale of the game, and therefore the information problem, increases.

As this is a problem inherent in wargaming, it manifests itself even in military institutions very different from those in the United States, and even in military institutions with comparatively more expertise in mathematical modeling and more access to granular data (Sloan 2020). Caffrey (2019) noted that Soviet military theorists struggled to create accurate wargaming models because even units at the same level of organization, such as battalions, tend to differ in terms of their allotted equipment and certainly in terms of their state of repair and training. Additionally, he noted that modeling a large wargame as simply the sum of multiple smaller wargames run with smaller units yielded different outcomes than the actual combined production observed in Soviet historical records of real conflict. Awareness of this problem caused the Soviets to comb through historical data looking for ways to determine the marginal effectiveness of different unit compositions to increase the fidelity of their models. For instance, using data from previous conflicts, they attempted to find the marginal change in construction time associated with assigning an additional engineer to an engineering battalion tasked with constructing bridges across rivers during combat. Their efforts increased the accuracy and predictive power of their models within the narrow confines of the singular unit type studied, provided no changes occurred in military technology or manning or training or resources between examples studied. However, once the scope of the games increased to include heterogeneous unit compositions, such as engineers with different tools or attached to mechanized infantry units, game results began failing to track with historical data. Moreover, once the scale of the games increased to include a larger number and variety of units, the information problems inherent in wargaming caused the games to depart even further from historical records (Postalov 1969).

Decentralized information is also held by the wargamer's enemies. This information is intentionally held secret by the wargamer's enemies and requires that the wargamer make educated guesses about enemy capabilities and intentions while playing as Red Forces. Given the strategic value of hidden information, there is no way to avoid this problem, but attempting to model Red behaviors from Blue priors can lead to self-deception. To illustrate, tanks emerged as a major innovation during World War One and this development led to search during the interwar period for tactics and weapons which could counter them. The United States Army conducted live wargames known as the Louisiana Maneuvers and the Carolina Maneuvers in 1941 to test the efficacy of their tank destroyer units against Red Force tanks (Denny 1990). Tank destroyers were used aggressively, despite their negligible armor, and produced outsized effects in both games with tank destroyers knocking out 91% of Red Force tanks during the games with no loss of tank destroyers. By the time of these games, the US Army's tank branch was only a year old and had not developed the same institutional expertise and experience of the German Army's tank branch which was ten years ahead of them. Without comparable experience, Red Forces could not accurately play as their future enemies. As such, tank destroyers deployed during World War Two to fight German tanks in Northern Africa routinely suffered large losses to better armored German tanks that operated more aggressively than the American tanks the tank destroyers had trained against. The tank destroyer's most successful engagement was at El Guettar, where they destroyed only 52% of the German tanks present and did so at a cost of 66% of the tank destroyers present. The information produced by the Louisiana and Carolina Maneuvers had given the Army a false sense of confidence in tank destroyers. This caused an overinvestment of irrecoverable resources into a highly specialized form of military capital, and time spent training in highly risky tactics that may not have been employed absent the false confidence generated by these wargames. Similar problems in modeling enemy behavior and capabilities plagued wargaming during other wars. During the Vietnam War, games were used to model the use of gradual escalation as a signaling mechanism. The sending and receiving of the signal was tested using players from the American military, but the intended target of the signal was the North Vietnamese. In reality, the North Vietnamese interpreted the signal not as a commitment to the conflict, but as a desire to avoid it which would play into the strength of their longer time horizons (Bakich 2014). The War on Terror provides another example. Army wargames run prior to the war hinted at the possibility of sectarian violence after a war (Appleget and Cameron 2015), but failed to reveal the enormous role improvised explosive devices would play because they lacked the knowledge of locally available resources that improvised explosive devices depend on (Smith 2011; Wood 2018). In both cases, the games created false confidence about enemy behavior and capabilities that led to the malinvestment of both physical and human capital.

Section 3.1.2. Problems of the Quantity of Information Necessary to Run Wargames

Wargames cannot economize on the amount of information they need to run or on the information they generate in a way that facilitates low-cost communication of their results. Money serves as a common measure of value in markets, but no similar common measure of value for war has been found in body counts, infrastructure destroyed, bombs dropped, or any other measure. Attempts to find and use a common measure have resulted in serious miscalculation, even in variables that can have prices attached to them such as value of life or infrastructure (Olson 1962; McMaster 1997; Wood 2018). Miscalculation occurs in part because of the omission of non-quantifiable variables such as morale, and in part because the leaders making military decisions often internalize only a negligible fraction of the cost of the war in terms of measures such as lives and infrastructure. The value of some action or some resource in war is typically determined by what the Austrian

Minister of the Department of War Economy during World War 1, Otto Neurath, called the 'direct consideration of the various possibilities' (Neurath 1919 p. 146). The difficulty in discovering all the relevant information held by friendly forces has been discussed already, but even assuming this knowledge is known the difficulty in consuming and transmitting it all remains high because there are so many various possibilities that require direct consideration. Most interventions and wars the United States has been a part of have lasted over a year and contained multiple theaters. A theater-level wargame lasting only three months can require between 200,000 to 300,000 pieces of data before even adding data on Blue force characteristics and Red force characteristics (Allen 1987), and the number of Blue and Red force operational units can number in the tens of thousands (United States Joint Forces Command 2002a).

Given that games are run by a multitude of branches and their subcommunities, the first step in facilitating the flow and consumption of the information these games generate would be a centralized repository of wargame designs and results. For decades in the United States, the need for such a repository was acknowledged but not addressed (Work and Selva 2015). This could be because wargames results do not generate enough reliable information to justify the cost of such a repository, it could be because individual branches or their subcommunities prefer to keep such information away from those they compete for a budget with, or it could be because the quantity of information such a repository would contain is too cumbersome to be of practical use. There is reason to doubt the reliability of the information generated by wargames, and incentive problems will be discussed later, but for now, the issue of the sheer quantity of information alone is enough to call into question the value of such a repository.

The idea that a repository of wargame designs and results should exist to facilitate the transmission of the information generated by these games goes back at least to the 1970s when computer simulation in wargaming was becoming more popular (Government Accountability Office 1971). GAO noted that the lack of a central repository resulted in duplication of wargaming efforts as well as a lack of both inter and intra-branch communication of design and results. Not only did this issue lead to scarce resources being wasted haphazardly reproducing data, but it also increased the quantity of information to be analyzed. The lack of a repository persisted for decades, and modern wargaming scholars were still publishing pieces on this problem less than a decade ago (Work and Selva 2015; Doubleday 2016; Pournelle 2016). A centralized repository of wargame results would be of value in mitigating the information problem wargaming faces, so the choice not to develop one suggests that that value is less than the costs of establishing such a repository. The results of the eventual establishment of a wargaming repository confirm that this was the case.

Heath and Svet (2018) record that the repository was developed in late 2015, funded with tens of millions of dollars, and made accessible to all military branches and all combatant commanders. In its first year, approximately 500 games and their results were loaded into the repository, but only 250 games between 2016 and 2017 were uploaded. Across the three years for which Heath and Svet provide data, 800 unique users accessed the repository, 216 of which were the officers who had run the wargame logging in to upload it. This figure represents less than half a percent of just the population of officers within the US military, never mind the military as a whole. Given that the repository essentially provides a public good in the form of free research on potentially life and death problems, the low level of engagement requires explanation. The unreliable nature of wargame results discussed in the preceding section offers one explanation that combines with another one. Members of the military are given an overwhelming number of tasks to perform. Wong and Gerras (2015) show that completing the total amount of training assigned to soldiers in the Army in a given year would require more training days than there are days in a year, leading to widespread deception about training completion. The opportunity cost of a serviceman's time is already very high, and wargame results tend to be hundreds of pages long (United States Joint Forces Command 2002b). Lincoln Bloomfield, who served as director for several wargames run by the Department of Defense, described this problem of communicating the information revealed by wargames the following way: ' ... and there is an absolute barrier right here ... so that much of it is not really passed along. People are so damn busy they can't read the morning paper, and you hand them a printout and they say, "Get this damn thing off my desk!"' (Allen 1987 pg. 158) The volume of information produced by wargames, even if this information is high quality and centralized in an easily accessible repository, is too great for overloaded military agents to engage with. This information problem insight answers the existing literature as to why a centralized repository was not developed earlier, but it also explains why that literature's calls for additional funding is unlikely to elevate wargaming from an intuition training device to a usable system of predictive modeling (Hernandez, McDonald, and Ouellet 2015; Work and Selva 2015; Appleget and Cameron 2015; Sloan 2020).

Section 3.2. Incentive Problems

The history of wargaming gives many examples of incentive problems causing real-world wargame performance to deviate from theoretical wargame performance. Agents at a variety of levels have an incentive to rig wargames, resulting in wargames producing information tailored to the interests of the agent rather than the honest results desired by their principals. Agents may seek to preserve weapons systems, reputations, or political priorities from the consequences of a wargame run honestly.

Weapons systems evolve, leaving professional communities built up around old systems in danger of becoming irrelevant and unfunded. These can include military communities such as Navy pilots or voters in districts that manufacture the weapons systems. The veteran, economist, and RAND analyst Jack Stockfisch described the incentive problem in wargaming this way:

The chance of getting honest numbers in this kind of setting, where the real war is between the Army, Navy, and Air Force over dollars, over budget (and if that war is going on within the Army, it's between the tankers and the artillery men and the infantry), the chances of getting detached honest analysis, of supporting empirical work in the field, of getting those field trials done honestly and in a hard-nosed way are very slim … But we don't want too much testing: we don't want hard information. Because when you get it too hard, the options available to decisionmakers are not available. If the evidence suggests clearly that you should get this type of airplane, or this type of tank, then you lose some degree of freedom with respect to where the contract goes, whether it's in Detroit, or Cleveland, or what have you. So you don't want it to be too rigorous … (Allen 1987, pg. 298)

With incentives to preserve budgets and status influencing agent behavior, wargaming cannot be said to be robust against manipulation. For instance, in 1981 the US Navy conducted a wargame titled Ocean Venture 81 wherein a naval force with two aircraft carriers tested its ability to sail across the Atlantic to the Norwegian Sea through waters contested by simulated Soviet naval forces, including submarines. According to the game's chief analyst, Lt Cmdr Dean Knuth, both carriers were sunk by simulated submarine attacks. This result was noted, and then the carriers were put back into play to complete other portions of the game. Knuth then submitted an article detailing the outcome of the game to the Navy's professional journal, *Proceedings*, where the Navy sat on it until the Senate Armed Services Committee inquired about the report. The Senate was in the middle of debating the appropriation of two more aircraft carriers with a total price tag of $6.3 billion, representing a significant acquisition for the Navy. To protect this acquisition, the day after the Senate Armed Services Committee's inquiry the Navy gave Knuth's article a secret classification which prevented it from being published and launched an investigation into the staff of *Proceedings* (Mintz 1982). This practice of obscuring the vulnerability of expensive and prestigious aircraft carriers turned into a persistent multi-decade norm across multiple U.S. Navy run wargames that was only broken when the inside story of Millennium Challenge 2002 was leaked (Wise 2015).

The information problems arising from decentralized information and the quantity of information provide ample opportunities for agents to cover the pursuit of their self-interest. During the wargame Nimble Dancer the Department of Defense (DoD) sought to test its ability to manage two major regional conflicts simultaneously (Government Accountability Office 1996a). United States Joint Forces Command (USJFCOM) sponsored the game and provided the Blue Forces, the Red Forces, and the White Cell to run the game. USJFCOM was also the organization tasked with developing the concepts being tested by the game, meaning no independent testing of the game occurred. The game was conducted and the Department of Defense confidently informed Congress that they were prepared to manage such a scenario should it arise. A follow-on investigation of Nimble Dancer conducted by the Government Accountability Office, however, revealed major flaws in the running of the wargame. USJFCOM officials conducted their wargame under the assumptions that Blue Forces would receive warning of enemy attacks in a timeframe early enough to make necessary adaptations, that the second major regional conflict would begin long enough after the first one to allow DoD time to establish a foothold and then split resources as the situation dictated, and that hostilities would cease at a time and in a way favorable to the redeployment of Blue Forces. GAO found that only limited testing of less favorable alternative timeframes occurred and that these less favorable timeframes were judged using different and more generous measures of effectiveness. Additionally, GAO found that while the USJFCOM reported that Nimble Dancer had demonstrated its ability to conduct two major regional conflicts even if chemical weapons were used, Red Force chemical weapon usage had been artificially constrained and testing of the impact of chemical weapons was contracted out to an independent firm using different measures of effectiveness than those used in the rest of the game. Finally, the effect of chemical weapons was tested in a war against a single regional power, not two. Not only did USJFCOM exclusively present the outcome of the game under its most favorable assumptions as proof that DoD was still meeting the requirement of being able to win two major regional conflicts, but it also relied on this tailored outcome to push for greater appropriations such as increased upgrades for the Air Force's bomber fleet. The need for these upgrades was disputed by GAO reporting that such upgrades were unnecessary given the availability of substitute weapon systems already available (Government Accountability Office 1996b).

A follow-on game to Nimble Dancer titled Dynamic Commitment was designed to further test DoD's ability to conduct two major regional conflicts, but instead demonstrated the persistence of the incentive problem. In this game, the Department of Defense claimed to have done due diligence in testing its ability to conduct two major regional conflicts against variations in DoD force levels. These were done at 10%, 20%, and 30% force level reductions, but these reductions were applied across all branches equally. GAO's investigation of the game revealed that these reductions and the uniform way in which they were applied were unnecessarily simplistic. More targeted reductions in certain weapons platforms or communities or the force levels of a particular branch would have preserved DoD's ability to fight and win two simultaneous major regional conflicts. This, however, would have localized budget and prestige cuts on particular branches or communities within those branches. Given that the wargame was co-produced by representatives from each military branch, any targeted reduction would have resulted in representatives withdrawing their support from the game. The intentionally simplistic, across-the-board reductions USFJCOM chose for the game helped them generate buy-in from each branch of the military and generated cover for favored platforms and communities by producing a result that allowed them to claim that tested reductions in DoD funding could threaten one of the core missions it was tasked with accomplishing (Government Accountability Office 1998).

Pressure to manipulate wargames sometimes comes from the political leaders who request that those wargames be run. In this case, the political leader seeks a certain outcome and influences either the rules by which the game is played or what data produced by the game makes it to the voting public serving as the politician's principal. Decades after the American Civil War had demonstrated the lethality of modern guns, the German Army began equipping more of its cavalry with lances for charging into gun lines. These charges were being reported as successful in German wargames as late as 1903, to the shock of American observers. When asked to explain why the game's umpires were adjudicating in favor of the cavalry charge instead of the gun line, they explained that they were pressured to manipulate the results of the game because of Emperor Willian II's enthusiasm for cavalry (Katzenbach 1984). Something similar happened with the Strategic Defense Initiative of the 1980s. The testing of this missile defense system was later admitted to be rigged to cause both Congress and the Soviet Union to overinvest in premature space-based defense technologies, with the post hoc rationalization being that this may cause the Soviet Union to bankrupt itself (Tirman 1993). Some of these tests intentionally used contractors with poor knowledge of Soviet military doctrine as Red Force players to make the missile defense system appear more effective than it actually could be (Allen 1987).

More often, however, pressure from political leaders to manipulate game results has to do with prioritizing important domestic policies over foreign affairs. During the Vietnam War, the Joint Chiefs of Staff sponsored a wargame called Sigma II-64 wherein the viability of the doctrine of gradually escalating the US military effort in response to North Vietnamese and Viet Cong activity was tested. The results predicted the actual US escalation as well as the actual response to it and the actual slow defeat and withdrawal that ensued. The Vietnam War was certainly a concern to President Johnson, but so were his election campaign and his proposed Great Society domestic programs. The more attention the failing war received, the more difficult it would be for Johnson to focus on pushing those domestic programs. Johnson repeatedly went to lengths to hide details of the war from both Congress and the public. For instance, he broke with the tradition of rotating chairmanship of the Joint Chiefs of Staff between the branches of the military to appoint General Wheeler as his Chairman. Johnson selected Wheeler for his sensitivity to the president's domestic political position and goals rather than his military expertise (Wheeler 1969), and Wheeler obliged by instructing ' the MACV commander-designate to coordinate his actions very closely with the embassy in Saigon and to avoid statements or actions potentially embarrassing to the Johnson administration' (McMaster 1997, pg. 110). Johnson's advisers also coached Congressional Democrats to 'downplay the seriousness of the situation in Vietnam by highlighting the "strengths of the South Vietnamese Army", which possessed "superior firepower and mobility"' (McMaster 1997, pg. 117). It is no surprise, then, that after Johnson's politically sensitive Secretary of Defense and Chairman of the Joint Chiefs of Staff personally observed the prophetic Sigma II-64 they intentionally ignored the results because neither escalation nor withdrawal was politically viable for the Johnson administration (McMaster 1997). Three earlier versions of the Sigma series of wargames had tested similar parameters, made similar predictions of gradual escalation producing a long war that ultimately favored the North Vietnamese, and were similarly ignored by the president and his advisors ((Ball 1992; Gibbons 1995; McNamara 1995). Game runners can manipulate the outcome of a game for their own interests, and the political leadership that orders games to be run can bury their results for their own interests.

The persistence of the incentive problem in wargaming shows that even if the knowledge problems were solved it would still be inadvisable to rely on wargame results for predictive purposes because the results will frequently be tailored to support outdated weapons systems, outdated doctrines, or unrelated domestic policies. This danger is mitigated the more wargames are constrained to being intuition training devices like chess, a finding that supports the wargaming literature that is critical of using wargaming for predictive purposes (Beuch 1987; Perla 2008; Pournelle 2016; Barney 2020).

Section 4. Conclusion

Wargames do not always fail because of information and incentive problems, but these problems are common, significant, and persistent across contexts. Wargames have had memorable success in producing valuable information. Proponents of predictive wargaming point to the U.S. Navy's success in predicting the behavior of the Japanese during World War II in the wargames that supported War Plan Orange. However, they neglect the details of this success. These games were played yearly for a period of decades, they focused the staff and students of the Naval War College on this one problem to the exclusion of others, they did not predict the development of kamikaze attacks and other suicidal defenses, and their success has not been repeated (Caffrey 2019). Others argue that sufficiently powerful computers using artificial intelligence may push wargaming to become more practical for predictive purposes, but at best this is a solution to only some of the quantity of information problem and none of the decentralized information and incentive problems.

Still, the War Plan Orange example offers an important lesson for those who wish to improve the outputs produced by wargames. Wargames are more likely to make accurate predictions if the scope of potential adversaries and theaters of operations is sufficiently limited, as it was to Japan and the Pacific, and a significant portion of analytical resources are devoted to studying conflict with those adversaries in those theaters. These limitations must be combined with many variations on the conflict being gamed, and probably tested across years or decades. Most of the wargames run this way will not be relevant to the conflict as it actually unfolds, but the sheer number of games run makes it likely that at least one will correspond to the real conflict and thereby offer insights into how to best fight it. Wargames are better focused and therefore more accurate when a nation's foreign policy is more limited because this reduces the quantity of information necessary to run relevant wargames. Focusing wargaming on a smaller number of potential adversaries has the additional benefit of making it more likely that rigging will be detected if conflict breaks out, decreasing an agent's incentive to rig in the first place.

This paper has argued that the information and incentive problems faced by wargaming are persistent. In light of this, the split in the current literature between those who argue for using wargames as models for forecasting and those who argue for using wargames to build the intuition of military decision-makers must be decided in favor of the latter. This means that games presented to the public as validating the performance of and need for some particular weapons system or operational concept should be viewed skeptically. It would be far more prudent to not allow wargames to be the justification for converting scarce resources into niche outputs that those resources cannot be easily recovered from. This would allow for greater adaptability to the far more reliable information revealed by the process of real-world conflict rather than simulated conflict.

Disclosure statement

No funding was received to assist with the preparation of this manuscript

ORCID

Garrett R. Wood http://orcid.org/0000-0001-6857-3735

References

Acemoglu, D., and J. Robinson. 2013. "Economics versus Politics: Pitfalls of Policy Advice." *Journal of Economic Perspectives* 27 (2): 173–192. doi:10.1257/jep.27.2.173.

Allen, T. 1987. *War Games: The Secret World of the Creators, Players, and Policy Makers Rehearsing World War III Today*. New York: McGraw-Hill Book Company. New York.

Anderson, G., and D. Dilegge 2015. Six Rules for Wargaming: The Lessons of Millennium Challenge '02. *War on the Rocks*. Available at: https://warontherocks.com/2015/11/six-rules-for-wargaming-the-lessons-of-millennium-challenge-02/

Appleget, J., and F. Cameron. 2015. "Analytic Wargaming on the Rise." *Phalanx* 48 (1): 28–32.

Bakich, S. 2014. *Success and Failure in Limited War: Information and Strategy in the Korean, Vietnam, Persian Gulf, and Iraq Wars*. Chicago: University of Chicago Press.

Ball, M. 1992. *Vietnam-on-the-Potomac. Praeger*. Westport: Connecticut.

Barney, R. 2020. "Thirty Degrees Off Axis." *Phalanx* 53 (2): 38–43.

Beuch, W. 1987. "Intelligence and Wargaming." *Phalanx* 20 (3): 10–12.

Brandes, S. 1997. *Warhogs.* Lexington: University Press of Kentucky.

Brennan, G., and J. Buchanan. 1985. *The Reason of Rules. Constitutional Political Economy*. Cambridge: Cambridge University Press.
Brown, T., and E. Paxson 1975. "A Retrospective Look at Some Strategy and Force Evaluation Games". *Rand*. Available at: https://www.rand.org/pubs/reports/R1619.html
Buchanan, J. 2003. "Public Choice: Politics Without Romance." *Policy* 19 (3): 13–18.
Caffrey, M. 2019. *On Wargaming*. US Naval War College. Newport: Road Island.
Clem, H. 1983. *Mobilization Preparedness*. Washington D.C: Nation Defense University.
Coyne, C. 2015. "Lobotomizing the Defense Brain." *The Review of Austrian Economics* 28 (4): 371–396. doi:10.1007/s11138-015-0316-x.
Coyne, C., and A. Hall. 2019. "State-Provided Defense as Noncomprehensive Planning." *Journal of Private Enterprise* 34 (1): 75–85.
Denny, B. 1990. The Evolution and Demise of U.S. Tank Destroyer Doctrine in the Second World War, [Master's Thesis, Appalachian State University]. Defense Technical Information Center. https://apps.dtic.mil/dtic/tr/fulltext/u2/a416377.pdf
Doubleday, J. 2016. "Navy Developing Integrated Wargame Plan to Better Share Lessons Learned." *Inside Missile Defense* 22 (7): 11.
Fearon, J. D. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379–414. doi:10.1017/S0020818300033324.
Gibbons, W. 1995. *The U.S. Government and the Vietnam War*. New Jersey: Princeton University Press. Princeton.
Government Accountability Office. 1971. Computer Simulation, War Gaming, and Contract Studies: Comptroller General's Report to the Committee on Appropriations House of Representatives. *National Security and International Affairs Division*. Available at: https://www.gao.gov/assets/120/112404.pdf
Government Accountability Office. 1996a. Bottom-Up Review: Analysis of DOD War Game to Test Key Assumptions. *National Security and International Affairs Division*. Available at: https://www.gao.gov/assets/nsiad-96-170.pdf
Government Accountability Office. 1996b. Air Force Bombers – Options to Retire or Restructure the Force Would Reduce Planned Spending. *National Security Analysis*. Available at: https://www.govinfo.gov/content/pkg/GAOREPORTS-NSIAD-96-192/html/GAOREPORTS-NSIAD-96-192.htm
Government Accountability Office. 1998. Quadrennial Defense Review: Opportunities to Improve the Next Review. *National Security and International Affairs Division*. Available at: https://www.gao.gov/assets/nsiad-98-155.pdf
Hayek, F. 1945. "The Use of Knowledge in Society." *American Economic Review* 35: 519–530.
Heath, G., and O. Svet. 2018. We Run Wargames Programs for the Joint Staff: Here's What We've Learned. *War on the Rocks*. Available at: https://mwi.usma.edu/run-wargames-programs-joint-staff-heres-weve-learned/
Hernandez, A., M. McDonald, and J. Ouellet. 2015. "Post Wargame Experimentation and Analysis: Re-Examining Executed Computer-Assisted Wargames for New Insights." *Military Operations Research* 20 (4): 19–37.
Katzenbach, E. 1984. "The Horse Cavalry in the Twenty First Century." In *National Defense University. The Art and Practice of Military Strategy*. Washington DC: NDU Press.
Kirzner, I. 1984. "Economic Planning and the Knowledge Problem." *Cato Journal* 4: 407–418.
Lavoie, D. 19852016. *National Economic Planning: What Is Left?* Arlington. VA: Mercatus Center.
McMaster, H. 1997. *Dereliction of Duty*. Harper Perennial: New York.
McNamara, R. 1995. *In Retrospect: The Tragedies and Lessons of Vietnam.* New York: Times Books.
Melman, S. 1985. *The Permanent War Economy: American Capitalism in Decline*. New York: Simon & Schuster.
Mintz, M. 1982. Article Critical of Carriers Stamped 'Secret' by Navy. *The Washington Post*. Available at: https://www.washingtonpost.com/archive/politics/1982/05/04/article-critical-of-carriers-stamped-secret-by-navy/08799f9d-e252-4b39-8083-68d37316edb7/
Neurath, O. 1919. *Empiricism and Sociology*. Boston: D. Riedel Publishing Company.
Olson, M. 1962. "The Economics of Target Selection for the Combined Bomber Offensive." *Journal of Royal United Service Institution* 107 (628): 308–314. doi:10.1080/03071846209428669.
Perla, P. 2008. "Casualties in Wargames and Casualties in War." *Phalanx* 41 (4): 19–20.
Postalov, A. 1969. "Modeling the Combat Operations of the Ground Forces." *Military Thought* 3: 9–16.
Pournelle, P. 2016. "Revitalizing the Wargaming Community." *Phalanx* 49 (3): 10–12.
Punjani, S. 2012. The Iron Triangle Manifested: U.S. Air Force Tanker Lease 2001-2005 Case Study. *National Defense University*. Available at: https://inss.ndu.edu/Media/News/Article/693827/the-iron-triangle-manifested-us-air-force-tanker-lease-20012005-case-study/
Sloan, J. 2020. *Soviet Style Wargaming*. Washington DC: Science Applications.
Smith, A. 2011. *Improvised Explosive Devices in Iraq, 2003-2009: A Case of Operational Surprise and Institutional Response*. Carlisle: Strategic Studies Institute.
Tirman, J. 1993. A Big Hoax to Match A Big Lie: Star Wars Was A Fantasy All Along, so Was the Soviet Threat. *Los Angeles Times*. Available at: https://www.latimes.com/archives/la-xpm-1993-08-20-me-25495-story.html
United States Joint Forces Command, 2002a. U.S. Joint Forces Command Millennium Challenge 2002: Experiment Report. *Joint Forces Command*. Available at: https://www.esd.whs.mil/Portals/54/Documents/FOID/Reading%20Room/Joint_Staff/12-F-0344-Millennium-Challenge-2002-Experiment-Report.pdf

United States Joint Forces Command. 2002b. United States Joint Forces Command Millennium Challenge 2002: Experiment Report. *United States Joint Forces Command*. Available at: https://www.esd.whs.mil/Portals/54/Documents/FOID/Reading%20Room/Joint_Staff/12-F-0344-Millennium-Challenge-2002-Experiment-Report.pdf

Wheeler, E. 1969. Oral History Transcript 1, 8/21/1969. Interview. Conducted by Dorothy Pierce

Wise, D. 2015. The U.S. Navy's Big Mistake – Building Tons of Supercarriers. *War is Boring*. Available at: https://medium.com/war-is-boring/the-u-s-navy-s-big-mistake-building-tons-of-supercarriers-79cb42029b8

Wong, L., and S. Gerras. 2015. *Lying to Ourselves: Dishonesty in the Army Profession*. Pennsylvania: United States Army War College Press.

Wood, G. 2018. "The Enemy Votes: Weapons Improvisation and Bargaining Failure." *Economics of Peace and Security Journal* 13 (1): 35–42. doi:10.15355/epsj.13.1.35.

Work, B., and P. Selva. 2015. Revitalizing Wargaming Is Necessary to Be Prepared for Future Wars. *War on the Rocks*. Available at: https://warontherocks.com/2015/12/revitalizing-wargaming-is-necessary-to-be-prepared-for-future-wars/

Zenko, M. 2015. Millennium Challenge: The Real Story of a Corrupted Military Exercise and Its Legacy. *War on the Rocks*. Available at: https://warontherocks.com/2015/11/millennium-challenge-the-real-story-of-a-corrupted-military-exercise-and-its-legacy/