Department of Economics

Allocating Security Expenditures under Knightian Uncertainty: an Info-Gap Approach*

Michael Ben-Gad[1]
City, University of London

Yakov Ben-Haim
Technion, Israel Institute of Technology

Dan Peled
University of Haifa

Department of Economics
Discussion Paper Series
No. 19/06

* This article is a substantial revised version of a working paper that was released in this series as DP No. 08/05.

[1] Corresponding author: Michael Ben-Gad, Department of Economics, City, University of London, Northampton Square, London EC1V 0HB, UK.
Email: mbengad@city.ac.uk.

Allocating Security Expenditures under Knightian Uncertainty: an Info-Gap Approach*

Michael Ben-Gad[a] Yakov Ben-Haim[b]

Dan Peled[c]

April 18, 2019

Abstract

We apply the information gap approach to resource allocation under Knightian (non-probabilistic) uncertainty in order to study how best to allocate public resources between competing defense measures. We demonstrate that when determining the level and composition of defense spending in an environment of extreme uncertainty *vis-a-vis* the likelihood of armed conflict and its outcomes, robust-satisficing expected utility will usually be preferable to expected utility maximisation. Moreover, our analysis suggests that in environments with unreliable information about threats to national security and their consequences, a desire for robustness to model misspecification in the decision making process will imply greater expenditure on certain types of defense measures at the expense of others. Our results also provide a positivist explanation of how governments seem to allocate security expenditures in practice.

JEL classification: D81; F51; F52

Keywords: Defense; Knightian Uncertainty; Robustness; Info-gap

*This research was supported by the Economics of National Security Program, Samuel Neaman Institute, grant no. 358. Ben-Gad: *mbengad@city.ac.uk*, Ben-Haim: *yakov@technion.ac.il*, Peled: *dpeled@econ.haifa.ac.il*

"*There are things we know that we know. There are known unknowns. That is to say there are things that we now know we don't know. But there are also unknown unknowns. There are things we do not know we don't know. So when we do the best we can and we pull all this information together, and we then say well that's basically what we see as the situation, that is really only the known knowns and the known unknowns. And each year, we discover a few more of those unknown unknowns.*[1]

Donald Rumsfeld

Former U.S. Secretary of Defense

1 Introduction

What is the best procedure for a policy maker to employ when allocating scarce resources to both deter military aggression, and in the event deterrence is unsuccessful, mitigate its effects, if previous experience provides little relevant guidance? This paper demonstrates how both the scope and nature of defense expenditure should change, if policy makers wish their decisions to be robust to the type of Knightian (non-probablistic) uncertainty they ordinarily confront.

When allocating resources, policy makers, like households, face trade-offs. Many of the trade-offs households encounter can reasonably be modeled in a deterministic setting—their incomes may be uncertain, but when deciding between a new refrigerator or a new television set, we generally assume that consumers know the marginal utility they will derive from each. By contrast, policy makers often confront decisions in which the connection between allocations, and the desirability of outcomes, are uncertain. Nonetheless they are often able to derive reliable probabilistic models that link different allocations of funding with the moments of a distribution of outcomes. For example, no one can predict precisely by how much a dollar transferred between different components of the public health budget will affect an individual citizen's longevity. However, policy makers know a great deal about not only the prevalence, but the distribution of infectious diseases, heart problems, and cancer for various population groups and the efficacy of different treatments for large samples of patients.

Unlike public health decisions, allocating resources for national security involves decisions where experiments are not possible (or at least unwise). Moreover, as relations between any set of political actors on the international stage are constantly shifting, and both military doctrines and technology continue to evolve, previous experience may not provide much useful information upon which to base present-day decisions on how best to cope with future threats.[2] The policy maker whose country is threatened by hostile forces may know little about the probability

[1] Antulio J. Echevaria II (2008).

[2] David Hume, writing in 1748: "What is the foundation of all conclusions from experience?....All experimental

distribution of possible damage or losses his or her country could suffer in the event of armed conflict, as any estimate will require both an intimate knowledge of the adversary's capabilities and tactics, and the degree to which they might be counteracted by different defense allocations. Indeed, the probability that an adversary will launch an attack may itself be unknown to the policy maker, as it requires judgements about the intentions and expectations of foreign leaders, or sub-state actors and yet is still dependent on the type of defense capabilities our policy maker has chosen. Finally, often enough the enemy's own political and military decision makers are uncertain about their own goals or the best means to achieve them, at the time our decision maker must determine the best allocation.

Many of the models designed to analyse military spending, starting with Richardson (1960), consider the strategic behaviour of two agents engaged in an arms race within a deterministic setting.[3] Niho and Takeuchi (2001) consider the optimal amount of deterrence achieved by overall defense spending as a deterministic resource allocation problem. Our work departs from the existing literature in two important ways. First, in place of the familiar 'guns versus butter' dichotomy between defense expenditure and consumption (Brito (1972) and van der Ploeg and de Zeeuw (1990), as well as Niho and Takeuchi (2001)), policy makers must also choose how best to spend the former in a world where both the level of overall defense spending and its internal allocation have uncertain and potentially differing effects on both deterrence and on the outcome of conflict, should deterrence fail. Furthermore, if the uncertainty surrounding either is non-probabilistic in nature—or as the evidence we present suggests, both are—merely shifting to a framework of resource allocation with expected utility maximisation will prove inadequate.[4] This is because the maximisation of expected utility demands that probability distributions are known, or at least that statistical moments can be calculated. Required is a methodology that enables the policy maker to allocate resources without requiring the use of unavailable probabilistic information. By implementing the info-gap methodology, policy makers combine what they know about security threats, and their relationships to installed military capabilities with specified policy requirements, without requiring the policy maker to know how wrong the available information is. Moreover, our analysis highlights one rationale for heightened spending

conclusions proceed upon the supposition that the future will be conformable to the past....why this experience should be extended to future times, and to other objects, which for aught we know, may be only in appearance similar; this is the main question on which I insist....If there be any suspicion that the course of nature may change, and that the past may be no rule for the future, all experience becomes useless, and can give rise to no inference or conclusion."(Hume 2004, pp. 19, 21, 22)

[3]Intrilligator and Brito (1981) introduce probabilistic uncertainty to model the impact of nuclear proliferation. The likelihood of an individual nuclear state initiating conflict declines as the number of nuclear states rises, creating a non-monotonic relationship between the outbreak of nuclear conflict and the number of nuclear-armed states.

[4]In Bier *et al.* (2007) a defender allocates resources to defend different targets to minimise expected losses. The adversary's preferred target is characterised by a known probability distribution.

on some types of defense measures when there is little reliable information about the random nature of threats to national security. This provides a possible positivist explanation for the way governments allocate these expenditures today, and how they have done so in the past.

In section 2 we describe the general formulation of the info-gap methodology in the context of defense expenditure. In section 3, by way of example, we consider a simple version of a dilemma faced by many policy makers today: how much to spend overall on defense and how much of that spending should be diverted towards the precision munitions, electronic warfare, and space-based communications and intelligence necessary to implement the 'Revolution in Military Affairs' (RMA) doctrine, and away from continued investment in large units employing traditional industrial-age military technology and platforms.[5] In section 4 we consider two historical examples. First, we analyse whether the decision by successive French governments during the interwar years to invest so heavily in the construction of the Maginot line represents a failure to allocate defense expenditure in a robust manner. Then we consider how the experience of US ground forces during the invasion of Iraq in 2003 demonstrated that many elements of the RMA associated "network-centric warfare" doctrine, introduced in the late 1990's to completely transform the US Army by 2025 (Farrell *et al.*, 2013), were profoundly non-robust to catastrophic failure.

Throughout this paper, the risk of being attacked and the probability distribution of damage conditioned on an attack occurring—both of which are info-gap uncertain—are related to both the size and structure of the military force policy makers have chosen. We demonstrate that when policy makers must operate in an environment in which information about the relationship between threats and different allocations for defense is unreliable, expected utility maximisation does not provide the best guidance for making decisions, because it is completely non-robust to model misspecification. Using a hypothetical specification of security risks and the cost effectiveness of alternative defense measures, we show that the desire of policy makers to attain some degree of robustness to model misspecification, will have profound effects on both the level of resources devoted to defense, and the distribution of these expenditures across the two possible types of defense measures.

[5]First elaborated during the 1980's by Soviet military theorists, in particular Marshal Nikolai Ogarkov, then chief of the Soviet General Staff, the RMA doctrine of war has already had a profound effect on Western military planning. The move to adapt the U.S. military to RMA type warfare was spearheaded by the Department of Defense's Office of Net Assessment, and the man who has headed it since 1973, Andrew Marshall, and the vice chairman of the Joint Chiefs of Staff between 1994 to 1996, Admiral William Owens. See Eliot A. Cohen (1996).

2 Formulation

Consider a country facing various threats to its national security. These threats can take various forms, including invasion by an aggressive neighbor or a terrorist attack. Policy makers perceive threats to security from all sources as a bivariate distribution that includes both the event of being attacked, and the damage the country will sustain conditional on the attack taking place. In contrast to standard stochastic optimization problems, the distribution itself is uncertain. Policy makers must decide what portion of the economy's resources they will devote to countering security risks, as well as how to allocate this expenditure between different defense measures, where each measure has a different effect on both the risk of an attack and the potential damage it will inflict. In this paper, this defense resource allocation dilemma is embedded within a standard economic framework in which the representative risk-averse individual in this country derives utility $u(c)$ only from consumption, c. The threats considered here, and the effect of any countermeasures, affect the resulting level of utility through their influence on the resources that remain for consumption, the likelihood of being attacked, and the conditional distribution of losses an attack may inflict.

Normalizing the economy's resources to 1, the policy maker must choose the fraction of all resources to devote to each of N different risk-mitigating expenditures $\chi = (\chi_1, \ldots, \chi_N)$. Without government debt, we require $\sum_i^N \chi_i \leq 1$ where $\chi_i \geq 0$ for all i. Any government expenditure detracts from the resources available for consumption, so that

$$c(\chi) = 1 - \sum_i^N \chi_i.$$

On the other hand these government expenditures reduce the likelihood of an attack and the fractional loss ψ in resources resulting from an attack if it takes place, where $0 \leq \psi \leq 1$.

The probability density function (pdf) of realised threats, conditioned on risk mitigating expenditures, is $p(\psi|\chi)$—a probability distribution very poorly known to policy makers. The best available estimate of $p(\psi|\chi)$ is denoted $\widetilde{p}(\psi|\chi)$, but it is incontrovertible that $\widetilde{p}(\psi|\chi)$ is highly unreliable. We denote the probability of attack, as a function of defense spending χ, by $P_{\mathrm{w}}(\chi)$, and the best (but highly uncertain) estimate of this function is $\widetilde{P}_{\mathrm{w}}(\chi)$. The salient feature of this function is that $\widetilde{P}_{\mathrm{w}}(\chi)$ is a decreasing function of each type of defense expenditure, holding expenditures on other types fixed.

Let $R(\chi|p, P_{\mathrm{w}})$ be the expected utility resulting from defense expenditure χ, when the probability of the attack being realised is P_{w}, and the pdf of the damage ψ is $p(\psi|\chi)$:

$$R(\chi|p, P_{\mathrm{w}}) \quad = \quad P_{\mathrm{w}}(\chi) \int_0^1 u\left[(1-\psi)c(\chi)\right] p(\psi|\chi)\, \mathrm{d}\psi + (1 - P_{\mathrm{w}}(\chi)) u\left(c(\chi)\right) \qquad (1)$$

where the last term $u(c(\chi))$ represents the level of utility agents in this economy enjoy if no security risks materialise.

The info-gap model in the current context is an unbounded family of nested sets of probability models of $p(\psi|\chi)$ and $P_{\rm w}(\chi)$, indexed by α, which represents the unknown horizon of uncertainty in the policy maker's best estimate of the chances of an attack occurring and of the conditional distribution of the damages it would inflict. We denote this info-gap model by $\mathcal{F}[\alpha, \widetilde{p}(\psi|\chi), \widetilde{P}_{\rm w}(\chi)]$, where $\alpha \geq 0$.[6] The important feature of this info-gap model is that larger α entails larger possible deviations of the actual distributions of threats from their best estimates, and accordingly a wider range of actual challenges the defense allocation chosen will need to confront. A specific illustration of the info-gap model will be presented in section 3.2.

The policy maker requires that the expected utility be no less than a critical value, R_c. This critical value may be uncertain, or the policy maker may wish to identify an acceptable value of R_c that can be confidently anticipated, given a proposed allocation χ. In other words, the goal of the policy maker is to maximise robustness to model uncertainty subject to a minimally acceptable expected utility. That is, the policy maker wishes to choose an allocation χ that yields the widest margin of error in model specification under which the expected utility equals or exceeds R_c.

Specifically, the robustness of allocation χ is the maximum horizon of uncertainty, α, up to which the expected utility is no less than R_c for all probability functions in $\mathcal{F}(\alpha, \widetilde{p}, \widetilde{P}_{\rm w})$:

$$\widehat{\alpha}(R_c, \chi) = \max\left\{\alpha \mid \ R(\chi|p, P_w) \geq R_c, \ \forall \ (p, P_w) \in \mathcal{F}(\alpha, \widetilde{p}, \widetilde{P}_{\rm w})\right\} \tag{4}$$

More robustness is better than less, so the robust-optimal allocation at the minimally acceptable reward R_c is the allocation that maximises the robustness:

$$\widehat{\chi}(R_c) = \arg\max_{\chi} \widehat{\alpha}(R_{\rm c}, \chi) \tag{5}$$

In summary, $\widehat{\alpha}(R_c, \chi)$ is the robustness of allocation χ for achieving expected utility no less than R_c. Likewise, $\alpha^*(R_c)$ below is the maximal robustness for achieving expected utility no less than R_c by choosing the appropriate values of χ. It is the maximum value of $\widehat{\alpha}(R_c, \chi)$ from

[6] Info-gap models obey two axioms:

(i) *Nesting* asserts that the range of possible pdfs increases as α increases:

$$\alpha < \alpha' \quad \Longrightarrow \quad \mathcal{F}[\alpha, \widetilde{p}, \widetilde{P}_{\rm w}] \subseteq \mathcal{F}[\alpha', \widetilde{p}, \widetilde{P}_{\rm w}] \tag{2}$$

(ii) *Contraction* asserts that when $\alpha = 0$, the best estimated models are the only possibilities:

$$\mathcal{F}[0, \widetilde{p}, \widetilde{P}_{\rm w}] = \left\{\widetilde{p}, \widetilde{P}_{\rm w}\right\} \tag{3}$$

These two axioms endow α with its meaning of a *horizon of uncertainty.* Henceforth we drop the explicit notation conditioning $\mathcal{F}$ on the defense allocation χ considered. It is, however, at the heart of the problem being analyzed in this paper that the stochastic nature of threats to the security of a country is influenced by the defense measures it undertakes, albeit in a manner that is not fully known to the decision makers.

(4), evaluated at R_c and $\widehat{\chi}(R_c)$ from (5).

$$\alpha^*(R_c) = \max_{\chi}\{\widehat{\alpha}(R_{\rm c}, \chi)\} = \widehat{\alpha}(R_c, \widehat{\chi}(R_c)) \tag{6}$$

As in any info-gap model (Ben-Haim, 2006), there is a fundamental trade-off between minimally acceptable reward and robustness to uncertainty: the former decreases as the latter increases as will be demonstrated subsequently.

It is useful to note the relationship between robustness and the smallest reward on $\mathcal{F}(\alpha, \widetilde{p}, \widetilde{P}_{\rm w})$. Let

$$M(\alpha, \chi) = \min_{(p, P_{\rm w}) \in \mathcal{F}(\alpha, \widetilde{p}, \widetilde{P}_{\rm w})} \{R(\chi | p, P_{\rm w})\} \tag{7}$$

It can then be shown that $M(\alpha, \chi) = R_c \Leftrightarrow \widehat{\alpha}(R_c, \chi) = \alpha$.

3 An Illustration with Two Types of Security Expenditure

3.1 Basic Structure

By way of example, suppose all military expenditures fall into one of two broad categories. First, there is the expenditure that incorporates recent innovations in military technology and tactics, based on the intensive use of information technology, high-precision weaponry, and satellites employed for both intelligence and command and control systems. We denote this type of defense expenditure, associated with the 'Revolution in Military Affairs' (RMA) doctrine, as χ_1. Second is the development and maintenance of large military formations, composed of large numbers of troops receiving traditional military training to serve as infantrymen or to operate armor, artillery, battleships and bombers. We denote this, the more traditional, industrial-age type of military expenditure, and the one most closely associated with twentieth-century warfare, as χ_2. Both χ_1 and χ_2 are measured as shares of GDP, and we assume there is no possibility of international borrowing. Recall that $p(\psi|\chi)$ is the probability density function (pdf) of the damage the nation sustains in terms of lost GDP in the event that it is attacked, and assume its shape is influenced by both the overall quantity of resources devoted to security, $\chi_1 + \chi_2$, and also by how resources are divided between χ_1 and χ_2. Next we illustrate the kinds of considerations which can be employed to set the shape of the probability models used in the sequel.

With its reliance on small, highly mobile and specially trained military units, the RMA doctrine can under the most optimal conditions limit the severity of casualties and other losses for many kinds of attacks, but is more prone to catastrophic failure, particularly in the event of a massive invasion by a large military force. Accordingly, we assume that holding the GDP share of total defense expenditures fixed, the more defense expenditures are weighted towards

RMA type, at the expense of traditional weapons systems and military formations, the lower are the expected losses conditioned on an attack occurring, but the risk of extremely high damages increases. By contrast, traditional large military formations can insure against the most severe and widespread losses, but reliance on them will entail higher losses under most circumstances. With all defense expenditure restricted to fall into these two broad categories, policy makers' best (but highly uncertain) estimate of the damage pdf, conditional on being attacked and given allocation of defense expenditure χ_1 and χ_2, is $\widetilde{p}(\psi|\chi_1, \chi_2)$. Given the long planning horizon necessary to prepare defense forces, the relevant unit of time in this model is a decade.

The functional form for the best (but nonetheless highly uncertain) estimate of the damage density function reflects all available knowledge about possible losses incurred in the event of suffering an attack, and how the risks of these losses relate to the values of χ_1 and χ_2. In the sequel we provide the specification of the particular probability density function we chose for our illustration. Here, it suffices to note the two most salient features implied by our specification, which capture the essence of the tradeoff assumed to exist between the two alternative types of defense expenditure:

1. Mean damage generally declines in each of the security expenditure types. That is, at least in the neighborhood of $\widehat{\chi}(R_c)$, we assume:
$$\frac{\partial \mathrm{E}(\psi)}{\partial \chi_i} < 0,\ i \in \{1, 2\} \tag{8}$$
2. At the same time, holding total security expenditure fixed, we expect to observe that increases in the share devoted to traditional security expenditures lower the probability of extreme damage, defined here as 60% of output. That is, we expect to generally observe:
$$\left.\frac{\partial \mathrm{Prob}(\psi < 0.6)}{\partial \chi_2}\right|_{\chi_1+\chi_2} > 0 \tag{9}$$

These damage estimates are conditional on the occurrence of an attack, but what do we know about its probability?[7] Clearly, determining the probability of attack will necessitate estimating the military capabilities of potential adversaries—capabilities they may varyingly wish to conceal or exaggerate.[8] More difficult still is the need to assess the intentions of foreign leaders, whose decision processes and perceptions are often informed by cultures and historical memories different from one's own.

[7] "As the ancient retiree from the Research Department of the British Foreign Office reputedly said, after serving from 1903-50: 'Year after year the worriers and fretters would come to me with awful predictions of the outbreak of war. I denied it each time. I was only wrong twice.'" (Hughes, 1976, p. 48)

[8] "An adversary can still decide to attack even though his capabilities are relatively weak (1) if he miscalculates the strength of the intended victim (as did the Germans in their attack on the Soviet Union in 1941, or the Arab States in their underestimation of Israeli capabilities in 1967); (2) if he is more interested in applying political pressure or making political gains even at the cost of military defeat; (3) if he gambles that his surprise attack will have a force multiplier effect sufficient to compensate for his inferior capabilities." (Handel, 1989, p. 241.)

> The assumption that all actors in the arena of world politics can be counted upon to take a rational approach in foreign policy is not without some validity and utility; but it does not offer a sufficient basis for estimating how these other actors view events, calculate their options, and make their choices of action....To describe behavior as "rational" is to say little more than that the actor attempts to choose a course of action that he hopes or expects to further his values. But, of course, what the opponent's values are and how they will affect his policymaking and decisions in different kinds of situations remains to be established. Moreover, foreign-policy issues are typically complex in that they raise multiple values and interests that cannot easily be reconciled. Even for the rational actor, therefore, choice is often difficult because he faces a value trade-off problem. How the opponent will resolve that dilemma is not easily foreseen, even by those close to him let alone by those in another country who are attempting to predict his action on the basis of the rationality of the assumption. (George, 1980, p. 66.)

Yet even if it were possible to understand a potential opponent's existing intentions in a manner sufficient to assign probabilities to an attack as a function of one's own defense allocations, this would still be insufficient. Intentions change.[9] Hence the probability of an attack over the course of a decade is subject to great uncertainty—policy makers can scarcely quantify with much confidence the probability of an attack long before an enemy might even be considering one.[10]

[9] Though the Japanese were not privy to German plans to attack the Soviet Union, during the months before Operation Barborossa was launched on 22 June, 1941, the Japanese actively considered beginning to prepare for a 'northern war'—abrogating the Neutrality Pact they had just signed with the Soviet Union on 13 April of that year (having fought unsuccessfully a brief war with it between May and September 1939), to launch an attack on that country's far eastern territories. Throughout that summer, civilians and military people in the Japanese government weighed the various merits of attacking the Soviet Union, attacking the Western powers to the south, or continuing to rely on diplomacy while consolidating gains in China and newly acquired French Indo-China. Only at a military conference on 3 September did the more aggressive view of the Japanese Army prevail over that of the more cautious Navy to initiate a war in the south, in response to continuing United States refusal to restore normal economic relations and not impede further Japanese ambitions in China. Three days later an imperial council approved the military's recommendation that "if, by the early part of October, there is still no prospect of being able to attain our demands, we shall immediately decide to open hostilities against the United States, Great Britain, and the Netherlands (see Butow (1961, p. 250).) The decision to begin preparations for the 'southern war' commenced only three months and a day prior to the attack on 7 December, 1941.

[10] "The Americans—and many Japanese—had thought that any attack on the distant American base at Pearl Harbor was impossible, given the inability of the Japanese fleet to conduct such distant operations due to refueling challenges and the need for radio silence along the long route to Hawaii." (Hansen, 2017, p.143)

3.2 A Fractional Info-Gap Model of Uncertainty

The density function $\widetilde{p}(\psi|\chi)$ is the best estimate of the pdf of damage from an attack, given security allocation $\chi = (\chi_1, \chi_2)$. However, this estimate is based on fragmentary and unreliable evidence, and hence contains potentially serious but unidentifiable errors. The same is true for the estimated probability of attack, $\widetilde{P}_{\mathrm{w}}(\chi)$. The true values deviate from these estimates by unknown amounts. We use a *fractional error* info-gap model to represent the gaps in both the pdf of the damage and the probability of attack (Ben-Haim, 2006). Let $\mathcal{P}$ denote the set of all pdfs on [0, 1]. For any $\alpha \geq 0$ our info-gap model consists of all density functions and probability values that differ proportionally from $\widetilde{p}(\psi|\chi)$ and $\widetilde{P}_{\mathrm{w}}(\chi)$, respectively, by no more than α. That is:

$$\mathcal{F}(\alpha, \widetilde{p}, \widetilde{P}_{\mathrm{w}}) = \Big\{ \, (p(\psi),\, P_{\mathrm{w}}) : \quad \begin{array}{l} p(\psi) \in \mathcal{P},\ |p(\psi) - \widetilde{p}(\psi|\chi)| \leq \alpha \widetilde{p}(\psi|\chi),\ \text{for all } \psi \\ 0 \leq P_{\mathrm{w}} \leq 1,\ |P_{\mathrm{w}} - \widetilde{P}_{\mathrm{w}}(\chi)| \leq \alpha \widetilde{P}_{\mathrm{w}}(\chi) \, \Big\}, \end{array} \tag{10}$$

The value of the fractional error α is unknown, so the info-gap model is not a single set, but rather an unbounded family of nested sets of possible pdfs and probabilities. Since α is unbounded from above, there is no known worst case.

3.3 Robustness Function

Conditioned on an attack occurring, the expected utility under the best estimates of damage density is:

$$\widetilde{r}(\chi) = \int_0^1 u\left[(1-\psi)c(\chi)\right] \widetilde{p}(\psi|\chi)\, \mathrm{d}\psi \tag{11}$$

where $c(\chi) = 1 - \chi_1 - \chi_2$

The unconditional expected utility, based on the best estimates of the damage pdf $\widetilde{p}(\psi|\chi)$ and the probability of attack $\widetilde{P}_{\mathrm{w}}(\chi)$, (as defined in (1)) is:

$$\widetilde{R}(\chi) = \widetilde{P}_{\mathrm{w}}(\chi)\widetilde{r}(\chi) + (1 - \widetilde{P}_{\mathrm{w}}(\chi))u(c(\chi)) \tag{12}$$

We assume that $\widetilde{r}(\chi) < u(c(\chi))$, so that the expected utility in the event of an attack is always smaller than it is in the absence of an attack. The expected utility for arbitrary $p(\psi|\chi)$ and P_{w} is obtained from (1). The objective of the policymaker, for a given R_c, is to choose the defense expenditures $\hat{\chi}(R_c)$ which attains an expected reward that does not fall below R_c for the widest horizon of uncertainty.

In order to find such a robust-optimal allocation we first need to express the robustness to model uncertainty ($\widehat{\alpha}$) as a function of the defense allocation. Full details of this representation are provided in the Appendix. Intuitively, we do that by exploiting the assumption that marginal utility is positive and decreasing in ex-post resources, and choose probability objects, to be denoted by $\hat{P}_w$ and $\hat{p}$, which are 'worse' than $\widetilde{P}_{\rm w}(\chi)$ and $\widetilde{p}(\psi|\chi)$, respectively, but differ from these proportionally by no more than α. Since an attack can only decrease expected welfare, the probability of attack is taken to be:

$$\hat{P}_w = \begin{cases} (1+\alpha)\widetilde{P}_{\rm w} & \text{if } \alpha < (1-\widetilde{P}_{\rm w})/\widetilde{P}_{\rm w} \\ 1 & \text{else} \end{cases} \tag{13}$$

For the damage pdf, we distinguish between $\alpha \leq 1$, and $\alpha > 1$. For the former we use:

$$\hat{p}(\psi|\chi) = \begin{cases} (1-\alpha)\widetilde{p}(\psi|\chi) & \text{if } \psi \leq \psi_{\rm m} \\ (1+\alpha)\widetilde{p}(\psi|\chi) & \text{else} \end{cases} \tag{14}$$

where $\psi_{\rm m}$ is the median of the best estimate of the damage pdf $\widetilde{p}(\psi|\chi)$.

For $\alpha > 1$ we use:

$$\hat{p}(\psi|\chi) = \begin{cases} 0 & \text{if } \psi \leq \psi_\alpha \\ (1+\alpha)\widetilde{p}(\psi|\chi) & \text{else} \end{cases} \tag{15}$$

where ψ_α satisfies:

$$(1+\alpha)\int_{\psi_\alpha}^{1} \widetilde{p}(\psi|\chi)\,\mathrm{d}\psi = 1 \tag{16}$$

In other words, ψ_α is the $1 - 1/(1+\alpha)$ quantile of $\widetilde{p}(\psi|\chi)$.

Next we define:

$$\widetilde{r}_1(\chi) = \int_0^{\psi_{\rm m}} u\left[(1-\psi)c(\chi)\right]\widetilde{p}(\psi|\chi)\,\mathrm{d}\psi \tag{17}$$

$$\widetilde{r}_2(\chi) = \int_{\psi_{\rm m}}^{1} u\left[(1-\psi)c(\chi)\right]\widetilde{p}(\psi|\chi)\,\mathrm{d}\psi \tag{18}$$

$$\widetilde{r}(\chi) = \widetilde{r}_1(\chi) + \widetilde{r}_2(\chi) \tag{19}$$

$$\delta_r(\chi) = \widetilde{r}_1(\chi) - \widetilde{r}_2(\chi) \tag{20}$$

$$\tag{21}$$

Since marginal utility is positive, $\widetilde{r}_1 > \widetilde{r}_2$. Consequently, as we show in Appendix A, the smallest expected utility $R(\chi|p, P_{\rm w})$ over $\mathcal{F}(\alpha, \widetilde{p}, \widetilde{P}_{\rm w})$ is obtained with $\hat{P}_w$ in (13), and $\hat{p}$ in (14) or (15) depending on whether the given R_c implies an α value less than or equal to 1 or greater than 1, respectively. The lowest reward corresponding to these probability objects is given by (25) and (27) in Appendix A, depending on whether α is smaller or greater than 1. Equating these lowest levels of expected rewards to R_c, and solving these non-linear equations for α yields the desired robustness as a function of the defense allocation chosen and the specified minimally acceptable reward R_c.

The value of $\widetilde{R}$ in (12) is expressed in terms of utility. In our economy the representative individual has a maximum of a single unit of consumption, from which defense expenditures and any damage to the economy are deducted. Define the allocation that maximises expected utility under the best estimates of an attack probability and damages pdf:

$$\chi^* \equiv \arg\max_{\chi} \tilde{R}\left(\chi|\widetilde{p}(\,\cdot\,|\chi), \widetilde{P}_{\rm w}(\chi)\right). \tag{22}$$

We use this allocation as a benchmark and define a loss function in terms of the compensating consumption differential q. For that purpose let $c(\chi^*) = 1 - \chi_1^* - \chi_2^*$ and define $L(q)$ to be the expected utility under $\widetilde{p}$ and $\tilde{P}_w$ obtained under defense allocation χ^* when an amount q of consumption is lost in addition to the amount $\sum_i \chi_i^*$ diverted to defense spending:

$$L\left(q\right) = \widetilde{P}_{\mathrm{w}}\left(\chi^*\right) \int_0^1 u\left[\left(1-\psi\right)\left(c(\chi^*) - q\right)\right] \widetilde{p}\left(\psi|\chi^*\right) d\psi + \left(1 - \widetilde{P}_{\mathrm{w}}\left(\chi^*\right)\right) u\left[c(\chi^*) - q\right] \tag{23}$$

Given any defence allocation χ, we find the unique value of q which satisfies $L(q) = \widetilde{R}(\chi)$, and denote this value by $\widetilde{q}(\chi)$. Since $L(q)$ is monotone in q, we can use $\widetilde{q}(\chi) = L^{-1}\left(\widetilde{R}(\chi)\right)$ to represent the welfare loss associated with allocation χ under $\widetilde{p}$ and $\widetilde{P}_{\mathrm{w}}$ compared to the expected utility associated with χ^*. Since by definition $\widetilde{R}(\chi) \leq \widetilde{R}(\chi^*)$, $\widetilde{q}(\chi) \geq 0$ for any χ. Likewise, the loss of utility associated with any minimally acceptable reward, $R_c \leq \widetilde{R}(\chi^*)$ is represented by its equivalent consumption loss $q_c = q(R_c) = L^{-1}(R_c)$. By definition, $L(0) = \widetilde{R}(\chi^*)$. The welfare loss associated with deviating from χ^* can be compensated for by a greater robustness to model uncertainty. This trade-off can be seen in Figure 1.

In Figure 1 the horizonal axis measures the sacrifice in expected utility expressed in consumption equivalent compared to the allocation χ^* for different defense allocations. The vertical axis measures the robustness to model uncertainty. Along each curve marked as χ^i we keep the defense allocation fixed, and display the trade-off between utility loss and robustness afforded by that allocation. Thus, the curve associated with χ^* begins at the origin, where the robustness is zero and so is the utility loss. Higher robustness can only be assured at lower minimally assured utility levels as shown by this upwards sloping curve. Other allocations provide different trade-offs along their curves. The *envelope* of all these curves gives the highest possible robustness for each utility loss on the horizontal axis, and depicts the allocation(s) that attain this combination of reward and robustness. For instance, expected utility loss no greater than q^0 is attained by allocation χ^0 under all probability models in $\mathcal{F}$ that differ fractionally from $(\widetilde{p}, \widetilde{P}_{\mathrm{w}})$ by no more than $\hat{\alpha}(L(q^0), \chi^0)$. Any other allocation that attains this maximal utility loss will do so at a *lower* robustness. Not every defense allocation is necessarily present on the envelope, as shown for example by allocation χ^2 in Figure 1. This schematic figure also shows that if the choice of defense allocations is restricted to a particular set, say $\{\chi^0, \chi^1\}$ whose curves intersect at q^1, then χ^0 provides higher robustness for all $q \leq q^1$ whereas χ^1 does it for $q > q^1$.

3.4 Numerical Results

This subsection presents the results of our analysis for particular specifications of the functional forms involved that were chosen for illustrative purposes. These specifications illustrate the intertwined channels through which defense allocations affect both the probability of suffering an attack and the losses such an attack would impose. Our detailed specifications are presented in Appendix B.

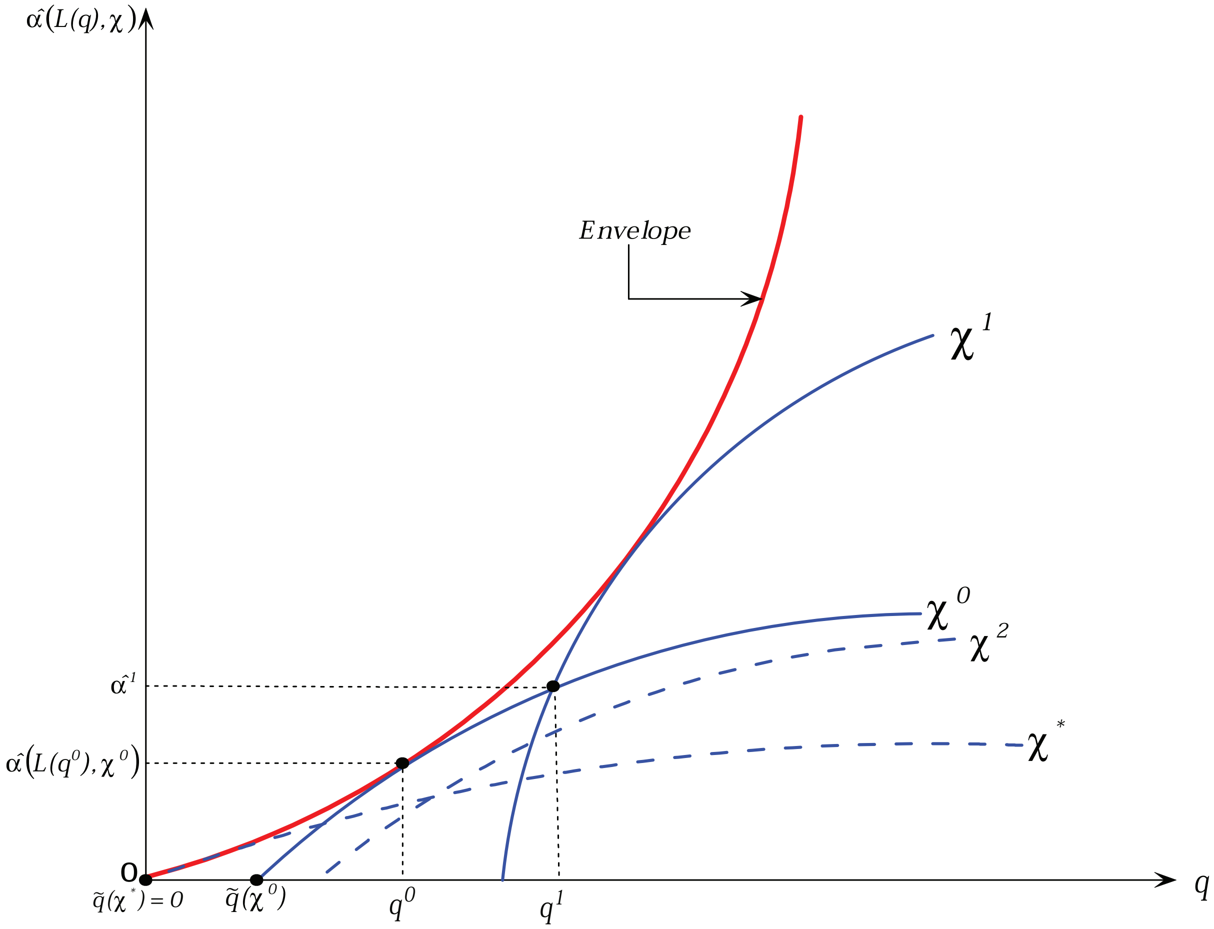

Figure 1: Each of the blue curves corresponds to a different allocation χ and captures the tradeoff between $\widetilde{q}$ and $\widehat{\alpha}$. The envelope in red is traced by the different blue curves and represents the maximum acceptable loss expressed as the consumption equivalent q_c and the maximum attainable robustness $\hat{\alpha}$ to model misspecification.

We focus on a scenario with relatively high values of potential damage in the event of attack, and choose parameters accordingly. For example, setting the overall size of the defense budget to equal 10% of total available income, equally divided between χ_1 and χ_2, implies given the functional specifications in (30) and (34), that expected losses in the event of attack will amount to an equivalent loss of 35.7% of annual consumption, where the best estimate of the probability of an attack taking place is 40.4%. Doubling the resources devoted to each, leaves the conditional loss equal to 32.3% and reduces the probability of attack to 30.8%, and doubling the resources yet again still leaves an expected loss of 26.4% conditional on being attacked with a best estimate attack probability of just under 19.3%.

A policy maker who seeks to *maximise expected welfare* under (30) and (34) will devote just under 6.38% of available resources to defense, with slightly more than half that, (χ_1=0.0427), allocated to RMA-type expenditure and the remainder, (χ_2=0.0211), to traditional military spending. Yet what is important to emphasise is that these probabilities of attack and damage density functions are merely best estimates, and the allocation that maximises expected utility is not robust to any degree of model uncertainty.

Each frame in Figure 2 shows plots of allocations χ_1 and χ_2 that yield constant robustness. Each frame in the figure is drawn for a different value of the lowest acceptable expected utility, R_c, expressed in equivalent consumption-loss, q_c = 0.05, 0.1, 0.15 and 0.2. We see that the maximum robustness values corresponding to these minimal expected utility levels are $\widehat{\alpha}$= 0.1442, 0.3016, 0.4746 and 0.6665, respectively. We note that greater robustness is attainable when greater expected consumption-loss is allowed. The contours in each frame represent the trade-offs between χ_1 and χ_2 for the fixed value of q_c of that frame, and the degree of robustness $\widehat{\alpha}$ is marked next to the contour. In Appendix A we derive the maximum robustness with which an allocation χ attains a minimally accepted reward R_c, ($\widehat{\alpha}$ in (4)), for the fractional info-gap model. We use these calculations to draw contour plots of constant allocations yeilding the same maximum robustness. We now explain the implications of this figure in greater detail.

Consider the first panel in Figure 2 which corresponds to the consumption equivalent loss of q_c=.05. In Appendix A we explain how we calculate $\widehat{\alpha}(R_c, \chi)$ – the maximum robustness of a minimally accepted reward R_c for a defense allocation χ – for the functional specifications described in Appendix B. Calculating the value of $\widehat{\alpha}$ for different combinations of χ_1 and χ_2 in (26) and (28) yields the contours, each one corresponding to a different value of $\widehat{\alpha}$ that surround the unique maximum at $\widehat{\alpha}$=0.1442. That maximum is attained at χ_1=0.0515 and χ_2=0.0402. This means that a policy maker willing to accept utility equivalent of 5% less consumption compared to that associated with expected utility maximisation can, by slightly increasing expenditure on RMA-type weaponry from χ_1=0.0427 to χ_1=0.0515, but more than doubling expenditure on traditional military hardware from χ_2=0.0211 to χ_2=0.0515, ensure at least 95% of the (consumption equivalent) expected utility maximising reward even if the best estimates of both the likelihood of attack, and the conditional damage are incorrect by as much as 14.42%.

The remaining three panels in Figure 2 further demonstrate the trade-offs between how much a policy maker is prepared to accept a lower level of utility (and hence higher values of q_c), and the maximum degrees of robustness to model uncertainty that this sacrifice affords. Thus by setting q_c=0.1, the degree of maximum robustness rises to 30.16%, achieved by setting χ_1=0.0597 and χ_2=0.0622; setting q_c=0.15, the degree of maximum robustness rises to 47.46%, achieved by setting χ_1=0.0678 and χ_2=0.0874; and setting q_c=0.2, the degree of maximum robustness rises to 66.65%, achieved by setting χ_1=0.0761 and χ_2=0.1156. In each case higher robustness is achieved by allocating more resources to defense spending. This is hardly surprising. What is being chosen here are allocations of defense expenditure designed to insure against model misspecification that is too optimistic. Yet the increases are *asymmetric*—relatively little of

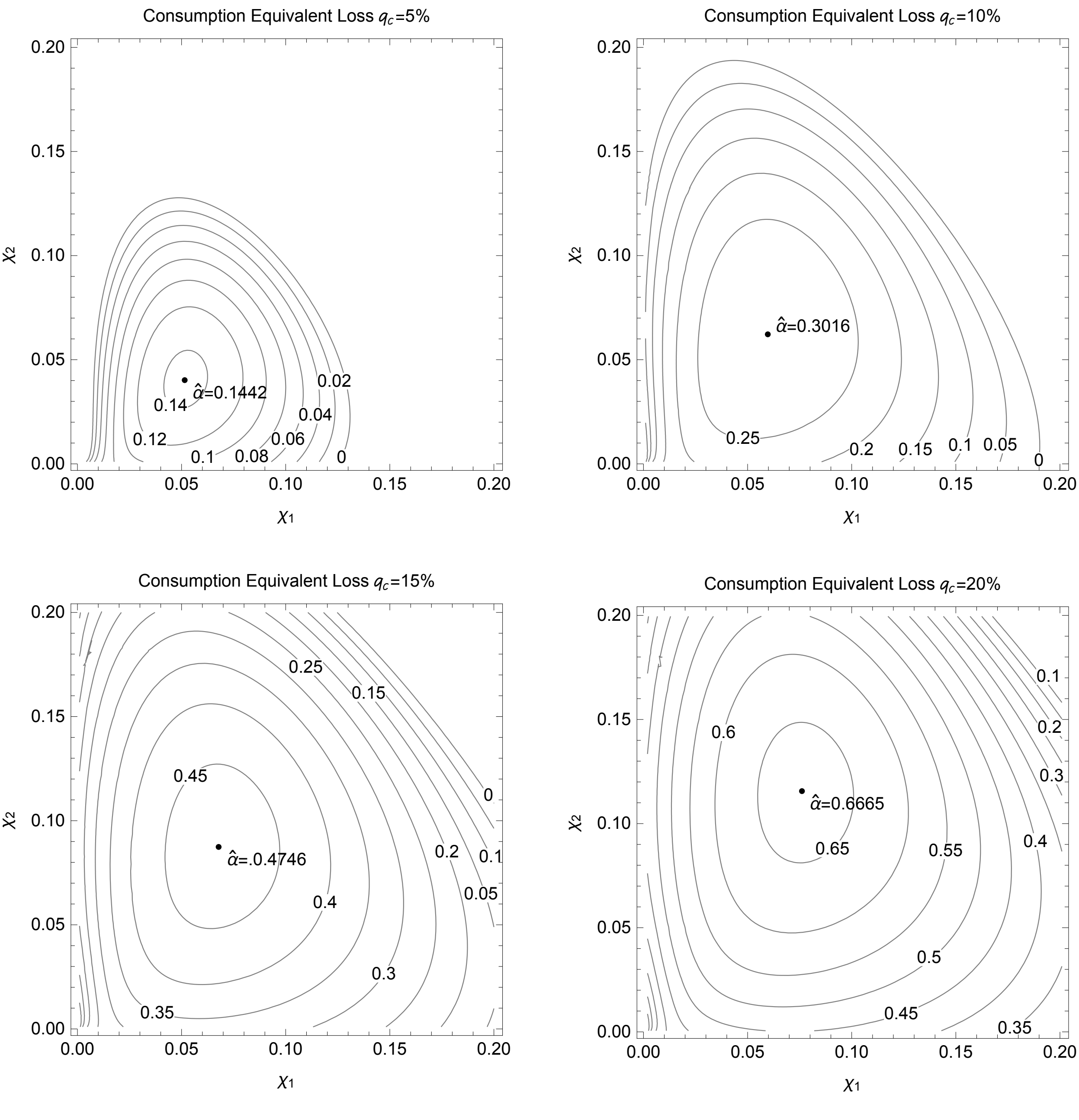

Figure 2: Contour plots of combinations of χ_1 and χ_2 generating constant robustness. Different contours correspond to different values of $\widehat{\alpha}$, and each panel is drawn for a different value of minimally acceptable value of expected utility – expressed in equivalent consumption loss terms $q_c = 5\%$, 10%, 15% or 20% relative to maximised expected utility.

the additional expenditure is allocated to the new RMA-type weaponry. Instead higher levels of robustness are mostly achieved by sharply increasing the investment in the more traditional weaponry.

Why should the share of resources devoted to traditional war fighting capacity rise so much faster than RMA-type expenditures as the desired degree of robustness to model uncertainty increases? The reason is that while increasing both types of security expenditure may generate higher levels of deterrence and also lower the amount of damage suffered in the event of an attack, it is the acquisition of planes, tanks and artillery, χ_2, rather than high precision weaponry and enhanced intelligence gathering, χ_1, that works best in lowering the likelihood of extreme damage.

Table 1 summarises the results in Figure 2. It emphasises the relationship that exists between maximum robustness, $\widehat{\alpha}$ and the acceptable level of losses q_c, and how they are achieved by different combinations of the two types of expenditure. The last column in Table 1 lists the values of $\widetilde{q}$ that express how much in consumption terms we actually expect to relinquish relative to expected utility maximisation under the best estimates $\widetilde{P}_{\mathrm{w}}$ and $\widetilde{p}$, (rather than q_c which represents the maximum reward we are prepared to relinquish). That is, while the maximum loss of expected utility will be smaller than q_c for *all* eventualities in a $\widehat{\alpha}$ surrounding of $\widetilde{P}_{\mathrm{w}}$ and $\widetilde{p}$ in any of the rows in Table 1, the expected loss of utility relative to maximised expected utility under the best available information is only $\widetilde{q}$.

It is important to emphasise the distinction between two different decision strategies. An expected utility maximiser allocates expenditure while paying little attention to low probability events associated with extreme losses. In contrast, a robust satisficer, wary of relying too heavily on uncertain estimated probabilities, will favour higher expenditure with the lion's share of additional resource devoted to spending that is less susceptible to catastrophic failure. This illustrates how the search for robustness alters decision making when the trade-offs satisfy (8) and (9). Nonetheless, the robust satisficer is not necessarily conservative, because the critical consumption loss, q_c, can be chosen at any value from cautious to cavalier.

Another way to understand the distinction between the expected utility maximiser and the robust satisficer is to compare the two expansion paths in Figure 3 (detailed in Table 2). The thick path to the right traces the allocations the former will choose (the tangencies between the iso-expected utility contours—fixed values of $\widetilde{q}$—and the dashed budget lines), while the arrows to the left trace the choices made by the latter, each for any given overall level of defense expenditure along its corresponding budget line. On any given budget line, the allocation chosen by the robust satisficer generates a lower level of $\widetilde{q}$, while the allocation chosen by the expected utility maximiser generates a lower level of robustness. This phenomenon is illustrated by each dashed iso-robustness contour, as it passes through the expected utility maximising allocation,

q_c	χ_1	χ_2	$\hat{\alpha}$	$\widetilde{q}$
0	0.0427	0.0211	0.	0.
0.05	0.0515	0.0402	0.1442	0.0033
0.1	0.0597	0.0622	0.3016	0.0126
0.15	0.0678	0.0874	0.4746	0.0277
0.2	0.0761	0.1156	0.6665	0.0488
0.25	0.0846	0.1468	0.8815	0.0757
0.3	0.0931	0.1806	1.1479	0.1082
0.35	0.1017	0.2111	1.4931	0.1411

Table 1: The combinations of χ_1 and χ_2 that generate maximum robustness $\widehat{\alpha}$, for different levels of maximally acceptable consumption equivalent losses compared to expected utility maximization, q_c, and the implied loss in expected utility (in consumption equivalents) of these defense allocations $\widetilde{q}$.

which is associated with a lower robustness level than the one provided by the robustness maximising allocation at the intersection of the same budget line with the robustness expansion curve (at the points designated by the values of $\widehat{\alpha}$).

To take one example, suppose the policy maker is faced with a fixed defense budget of 20% of GDP. As can be seen in Table 2, an expected utility maximiser will choose the allocation χ_1=0.0921 and χ_2=0.1079. By contrast, the robust satisficer, setting q_c=0.2108, will raise the spending on χ_2 by 13.2% to χ_2=0.1221 (and lower χ_1 to 0.0779), and by doing so achieves robustness of 71.08% while the expected utility loss in consumption equivalent terms increases only from 0.0527 to 0.0541.

To better understand the implicit tradeoff between robustness and minimum acceptable levels of utility, we plot in Figure 4 the implicit relationship between $\widehat{\alpha}$ and q_c for combinations of χ_1 and χ_2 between 0.01 and 0.2 (in increments of 0.01).[11] Each of the 400 grey curves represents a particular allocation of defense expenditure (χ_1, χ_2) and traces the implicit relationships in (26) and (28) between different values of $\widehat{\alpha}$ and q_c (with the aid of the loss function (23)). Each curve corresponds to a particular defense allocation, and begins on the horizontal axis where $\widehat{\alpha}$=0 and the critical welfare loss, q_c, equals the estimated value, $\widetilde{q}$ evaluated under $\widetilde{p}$ and $\widetilde{P}_{\mathrm{w}}$. As we consider larger deviations from the best guess uncertainty model, (i.e. increase $\widehat{\alpha}$), that particular allocation yields larger welfare losses, and hence the curves are upward sloping. The upper contours of all the grey curves generate the *envelope* that traces the relationship between maximum robustness $\widehat{\alpha}$ and q_c, starting at the origin which corresponds to the policy that generates maximum expected utility and zero robustness.

Policy makers can first determine the minimum level of acceptable utility, denominated here

[11] This figure maps the specific functional forms and parameter values used here into the analogue of Figure 1

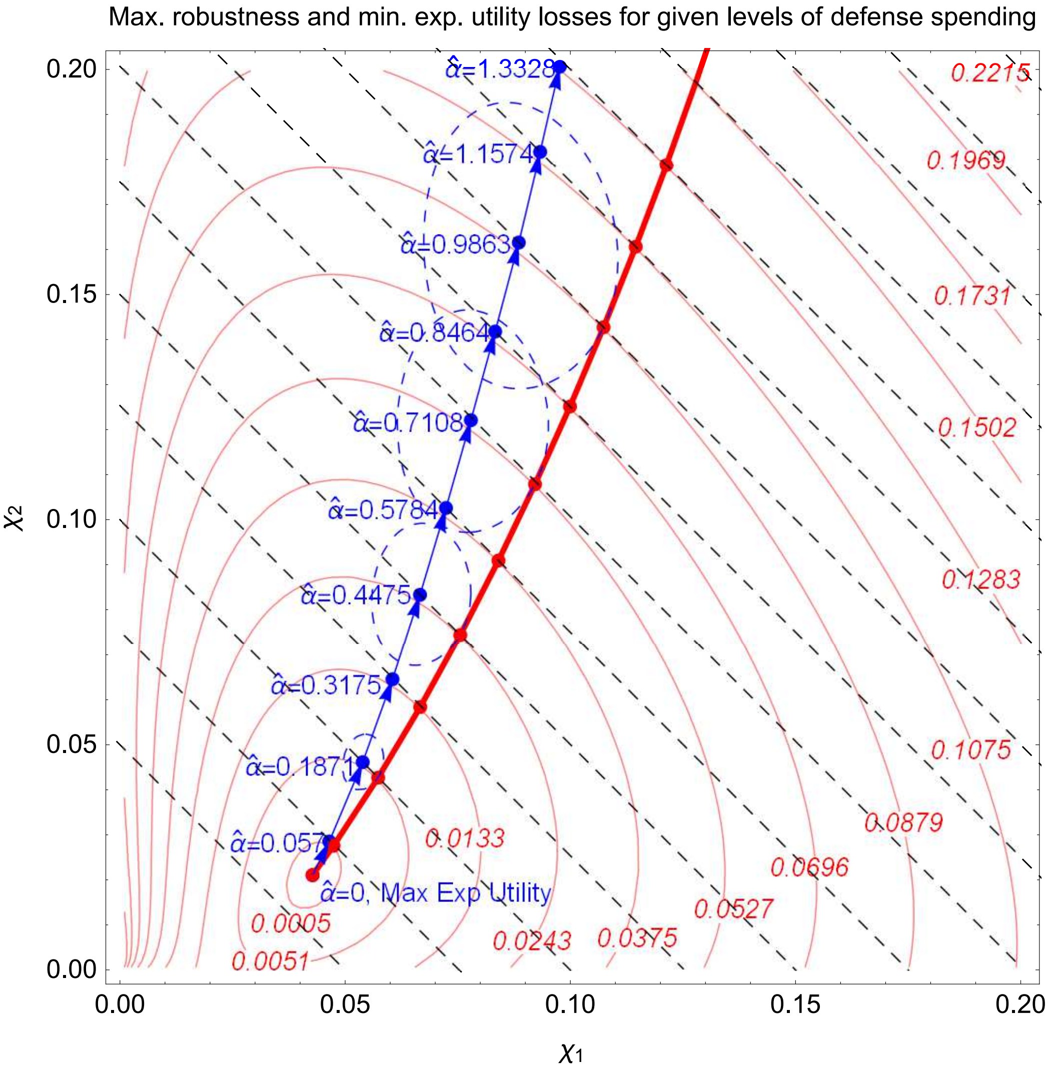

Figure 3: Dashed straight lines represent budget lines for given levels of defense spending (first column in Table 2). The solid thick curve represents all the combinations of χ_1 and χ_2 attaining maximum expected utility for a given level of defense expenditure, (columns 2 and 3 in Table 2). The thin line with arrows depicts, for the same total defense spending, all combinations of χ_1 and χ_2 attaining maximum robustness $\widehat{\alpha}$ (last column in Table 2) for the minimally accepted loss q_c, (fifth column in Table 2). The solid contours represent the allocations that yield the same expected utility ($\widetilde{q}$) level as the allocation that maximises expected utility for the given total defense spending. Dashed contours depict all the allocations that assure a maximal expected utility loss of of q_c with the same robustness level attained by the expected utility maximising allocation corresponding to the same defense budget line.

Total Spending	Welfare-Loss Minimization			Robust Allocation				
$\chi_1 + \chi_2$	χ_1	χ_2	$\tilde{q}$	q_c	χ_1	χ_2	$\tilde{q}$	$\hat{\alpha}$
0.0638	0.0427	0.0211	0	0	0.0427	0.0211	0	0
0.075	0.0474	0.0276	0.0005	0.0203	0.0464	0.0286	0.0006	0.0570
0.1	0.0573	0.0427	0.0051	0.0641	0.0539	0.0461	0.0053	0.1871
0.125	0.0666	0.0584	0.0133	0.1048	0.0605	0.0645	0.0138	0.3175
0.15	0.0755	0.0745	0.0243	0.1425	0.0666	0.0834	0.0251	0.4475
0.175	0.084	0.091	0.0375	0.1777	0.0723	0.1027	0.0387	0.5784
0.2	0.0921	0.1079	0.0527	0.2108	0.0779	0.1221	0.0541	0.7108
0.225	0.0999	0.1251	0.0696	0.2422	0.0832	0.1418	0.0711	0.8464
0.25	0.1073	0.1427	0.0879	0.2723	0.0885	0.1615	0.0896	0.9863
0.275	0.1144	0.1606	0.1075	0.3015	0.0933	0.1817	0.1093	1.1574
0.3	0.1213	0.1787	0.1283	0.3300	0.0980	0.2020	0.1302	1.3468

Table 2: Welfare-loss minimization vs Robustness maximisation. For a given total defense spending in the first column, columns 2-3 display the allocations that maximise expected utlity (under best estimates of $\widetilde{p}$ and $\widetilde{P}_{\mathrm{w}}$), and the expected utility loss relative to the expected utility maximizing allocation in the first row, ($\widetilde{q}(\chi)$). Columns 4-8 display for the same total defense spending, the robustness maximising allocations for the minimally acceptable loss q_c in the fifth column. The expected utility loss is displayed in column 8, and the resulting robustness in column 9.

as consumption equivalent deviations from maximum expected utility, and this determines via the envelope curve in Figure 4 the maximum degree of attainable robustness to model error. A given combination of defense expenditure (χ_1, χ_2) delivers a monotonic relationship between the value of q_c and $\widehat{\alpha}$ in Figure 4, which together generate an envelope that defines the relationship between q_c and $\widehat{\alpha}$ (the additional blue curves correspond to the particular values of χ_1 and χ_2 that generate maximum robustness for the values $q_c = 0.05, 0.1, 0.15, 0.2$). The more policy makers are cognizant of the limitations of the information they possess, in terms of efficacy of the measures they have at their disposal to both deter and repel aggression, the more we should expect them to reach for solutions that achieve robustness through both greater military expenditure and heavier reliance on the type of expenditure that is most effective at preventing catastrophic losses.

4 Discussion

Often enough, the pursuit of policies designed to limit casualties without being robust to misspecification can prove, with hindsight, misguided. At other times, such policies can result in complete catastrophe. Soon after the end of World War I the French government and military immediately began a debate on the best way to prepare for renewed conflict with Germany. One doctrine relied heavily on a smaller professional army well-trained in offense, and capable of mastering the new techniques of maneuver. The alternative that was eventually adopted relied on a larger force of conscripts backed by reservists to defend France's borders along a static heavily fortified front. On 30 September 1927, the *Commission d'Organization des Région Fortifiées*, (CORF) was established to plan construction of a line of fortifications along France's western frontiers with Italy, Germany, portions of southern and central Belgium as well as Corsica, ultimately came to bear the name of the Minister of War and decorated veteran of Verdun, André Maginot. The Maginot line was meant to protect France from a possible German invasion once France had completed the planned evacuation of its forces from the Rhineland in 1935. The line was based on a series of concrete multi-storied casements and *ouvrages* (fortresses). The smallest housed a squad of thirty men manning twin 7.5mm machine guns and 37 or 47mm anti-tank guns. The largest, the *gros ouvrages*, were large complex structures, built mostly underground, housing garrisons of 500-1,000 men from where heavy artillery could be fired from turrets set in the ground. Each was surrounded by their own series of machine gun turrets for protection.

The Maginot line's appeal was its promise to both deter a German attack and, in the event deterrence failed, to minimise French casualties by replacing 'a wall of chests' with 'a wall of concrete'. Between 1930 and 1935, by which time most of it had been completed, its construction alone consumed between one-half and one percent of French output each year, at a time when total French expenditure on defense averaged 4.7% of GDP (Thies, 1987). Two important gaps

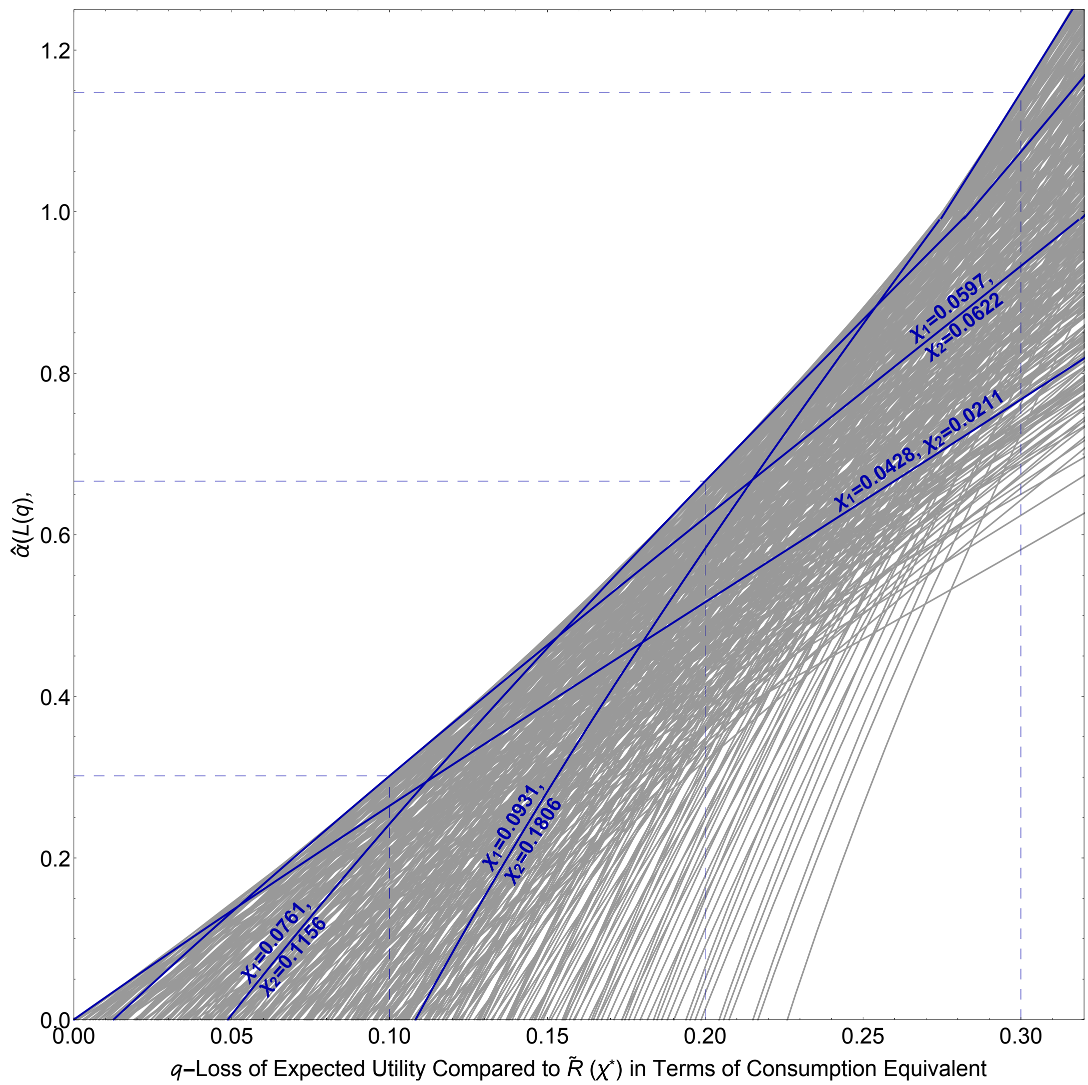

Figure 4: Each of the grey curves corresponds to a different combination of (χ_1, χ_2) and captures the tradeoff between $\widetilde{q}$ and $\widehat{\alpha}$. The envelope traced by the different grey curves represents the maximum acceptable loss expressed as the consumption equivalent q_c and the maximum attainable robustness $\hat{\alpha}$ to model misspecification.

remained along the border with Belgium: near the sea, where the high water table precluded the building of subterranean forts; and in the Ardennes, where Marshal Pétain, inspector general of the army until 1931, had asserted the terrain was impassible to German armour (Kaufmann and Kaufmann, 1997)

During the six-week German assault on France between 10 May 1940 and the signing of the armistice on 22 June, nearly every fort along the Maginot line performed as it was designed—they deterred direct attacks along the Franco-German border, and most successfully repulsed those efforts made by the Germans to attack them from the east. Ultimately, however, Pétain's assurances notwithstanding, the Germans did move their armour through the Ardennes, outflanking the Maginot line and forcing French capitulation. French strategic doctrine, designed to operate under one set of conditions, was not robust to the unexpected, and ultimately failed in a catastrophic manner. A doctrine that might have proven more robust to catastrophic failure—that advocated by General Charles De Gaulle (but successfully implemented by *Wehrmacht* General Heinz Guderian) of armoured forces able to move quickly to attack and fight independently of the infantry—was rejected, in part because it implied the likelihood of higher casualties (Kaufmann and Kaufmann, 2006)

In our interpretation, investment in the Maginot line was then akin to choosing high values of χ_1 (RMA in our more modern example). Had the Germans not found a way to circumvent it, it might have repelled the invasion, and done so as designed—at the cost of relatively few casualties. Instead it failed catastrophically—French planning was not sufficiently robust.

By contrast to the French experience during the second world war, America's successful invasion of Iraq in 2003 (Operation Iraqi Freedom) might seem in many regards vindication of the US military's embrace of the Revolution in Military Affairs doctrine. US forces toppled Saddam Hussein's regime in three weeks (between 20 March and 10 April 2003), at the cost of 139 American lives. Overall, less than one in every 2,300 coalition troops that participated in the invasion were killed in action (Biddle, 2007). In the initial aftermath much of the success was attributed to the adoption of high precision weaponry, and the Pentagon's embrace of the RMA-inspired "network-centric warfare" doctrine (i.e., "taking advantage of information technology to radically enhance the effectiveness of "C^4ISR"—command, control, communications, computers, intelligence, surveillance, and reconnaissance"), (Boot, 2003, p. 51).

The adoption of "network-centric warfare" was not merely a conventional upgrade of existing capabilities. The goal was to make the military, and in particular the US Army, quickly deployable by shifting from a reliance on heavy tracked armour (64 metric ton M1 Abrams tanks and 23 metric ton M3 Bradley armoured personel carriers) to forces mounted on wheeled lightly-armoured vehicles (17 metric ton Stryker troop carriers) that could be easily deployed by cargo planes (C-130 aircraft), but that lacked the armour necessary to protect soldiers against munitions more powerful than .50 calibre rounds. Technology and intelligence, combined with long-range, over-the-horizon, non-line-of-sight (NLOS) weapons designed to destroy the enemy from a distance, would replace armour by obviating the need to withstand close combat with the enemy (Kagan, 2007) while minimising civilian and military casualties (Futter, 2015). As stated by Stuart Johnson, National Defense University: "...information technologies allow you to substitute information for mass. If you buy into that, the whole force structure changes."…. And he continued, "But the vision of all this is totally dependent on information technologies and the network. If that part of the equation breaks down, what you have are small, less capable battle platforms that are more vulnerable,"(Talbot, 2004, p. 40).

Indeed as with the Maginot line, the new army, transformed according to the specifications of the RMA doctrine, could prove vulnerable to catastrophic failure, as it relies on the many elements of "network-centric warfare" to work perfectly. Kagan highlights the inherent non-robustness of a force overreliant on RMA:

> The Army's transformation program, in fact, came to rely even more heavily than the airpower visionaries on acquiring and disseminating perfect intelligence. The airpower enthusiasts needed this intelligence to target the enemy systems effectively, but Army units would require it merely to survive on the battlefield. (Kagan, 2007, p. 251)

The year 2003 found the US military only part way through adoption of this new doctrine, and the invasion of Iraq offered an opportunity to examine how those new systems, associated with the "network-centric warfare" doctrine which had already been introduced performed (our χ_1), in conjunction with the yet to be transformed (and heavily armoured) divisions that the US deployed (our χ_2). How easily, and at what cost in casualties, might the latter have executed the invasion had they undergone the full RMA transformation?[12]

New technologies did provide senior commanders unprecedented access to real-time intelligence and this made the air war particularly effective. The Blue force tracking system, designed to track friendly 'blue forces', radically reduced the number of casualties from 'friendly fire' incidents. Yet, in the years that followed, more nuanced interpretations of the events of 2003 emerged, suggesting that the success of the campaign at such a relatively low cost in lives of US and coalition personel could be largely attributed to the poor state and leadership of the Iraqi army rather than RMA. Beyond that, it became clear that particular systems at the very heart of "network-centric warfare" doctrine, those designed to pass along to frontline units on the ground the abundant real-time intelligence about the enemy 'Red force' that was streaming into command headquarters in Kuwait and Qatar, nearly all failed.

> The U.S. forces were all networked together, with a "blue force tracker" letting them know the position of all the friendly units, just as the network theorists had claimed would revolutionise war. The only problem is that they still didn't know who the enemy was ("the red force") or when he was coming. As one report dryly put it, "Situational awareness was proving to be more theoretical than actual." Or, as a marine joke, "When do we get red force trackers?"(Singer, 2009, pp. 189-190)

Mostly, this was an outcome of the Army's Mobile Subscriber Equipment (MSE) failure to support communications:

[12] The units that had already undergone the full transformation were either not deployed or had little influence on any part of the campaign (Adams, 2006).

> In order to employ the system, the division's signal battalion had to stop and set up their equipment. As soon as they did, the rapidly advancing units promptly ourtran the system's 15-mile radius....
>
> One brigade of the 3rd ID reported that whenever the unit moved, everything would fail except for the Blue Force tracking system. Every few hours the unit would stop, hoist up its antennas, log back into the network, and attempt to download whatever it could. But software and bandwidth problems would lock up the computer for 10 to 12 hours at a time, rendering it useless. Perversely, in three cases, U.S. vehicles were attacked while they stopped to receive intelligence on enemy positions. (Adams, 2006, p. 147)

Or stated more laconically in the Third Infantry Division's After Action Report (2003, p.4): "The communications available to all battlefield operating systems (BOS) other than maneuver were insufficient to ensure timely, accurate, and relevant information dissemination across the entire battlefield."

The 'digital divide' between the flow of information available to senior commanders and the dearth of information that flowed to frontline troops was exemplified by the experience of Lt. Col. Ernest Marcone, a Battalion commander in the 69th Armoured Regiment, during an operation code-named Objective Peach. Tasked with capturing the highly strategic *al-Kaed* bridge spanning the River Euphrates thirty kilometers southwest of Baghdad on 2 April, 2003, his unit was repeatedly ambushed on the way to its objective. The technology meant to provide Marcone with real-time intelligence failed completely: "I would argue that I was the intelligence-gathering device for my higher headquarters....Next to the fall of Baghdad that bridge was the most important piece of terrain in the theater, and no one can tell me what's defending it. Not how many troops, what units, what tanks, anything. There is zero information getting to me. Someone may have known above me, but the information didn't get to me on the ground."(Talbot, 2004, p. 38)

For Marcone's unit worse was to follow after the bridge was secured:

> As night fell, the situation grew threatening....One communications intercept did reach him: a single Iraqi brigade was moving south from the airport. But Marcone says no sensors, no network, conveyed the far more dangerous reality, which confronted him....He faced not one brigade but three: between 25 and 30 tanks, plus 70 to 80 armored personnel carriers, artillery, and between 5,000 and 10,000 Iraqi soldiers coming from three directions. This mass of firepower and soldiers attacked a U.S. force of 1,000 soldiers supported by just 30 tanks and 14 Bradley fighting vehicles. The Iraqi deployment was just the kind of conventional, massed force that's

> easiest to detect. Yet “We got nothing until they slammed into us“....Marcone only learned what he was facing when the shooting began. (Talbot, 2004, pp. 38, 44)

What proved decisive was the yet-to-be phased-out heavy armour, which unlike a completely transformed force, could operate effectively even when the larger systems of communications and intelligence failed:

> ...it was old-fashioned training, better firepower, superior equipment, air support, and enemy incompetence that led to a lopsided victory for the U.S. troops....Whereas U.S. tanks could withstand a direct hit from Iraqi shells, Iraqi vehicles would “go up like a Roman candle” when struck by U.S. shells....At Objective Peach, what protected Marcone’s men wasn’t information armor, but armor itself. (Talbot, 2004, p. 44)

Considering the likely outcome of Objective Peach had the tranformation of the army been completed and heavy armour not been available to Marcone, John Gordon of the Rand Corporation conjectured, “If the army had had Strykers at the front of the column, lots of guys would have been killed”, (Talbot, 2004, p. 44).

The original plan developed during the late 1990’s, to transform the US Army by 2025 and eventually replace entirely the Army’s heavy divisions, was indeed altered drastically as a consequence of the Iraq war. Farrell *et al.* (2013, p. 231) cite “the physical frailty of the current technology and the limits of what might be delivered in terms of enhanced battlefield knowledge,” along with vulnerability of the Stryker vehicles deployed in the aftermath of the invasion to improvised explosive devices and rocket-propelled grenades employed by the insurgents, for the cancellation of the US Army’s more ambitious transformation programmes.[13]

Of course nothing in our analysis suggests that military planners should eschew new technologies, even those technologies that are by their very nature vulnerable to catastrophic failure. What the info-gap approach does imply is that in an environment of non-probabilistic uncertainty, military planners do need to take into account this type of vulnerability and its potentially dire consequences when allocating limited defense resources. Kagan (2007, p. 352) cites “the failures of the network” as a reason for not eliminating heavy armour from any future force. As Biddle (2007, p.7) writes: “Instead, a balanced military in which the US can provide standoff precision fires *and* conventional, close-combat ground maneuver, en masse if necessary, continues to provide a valuable hedge against uncertainty.”

[13] As of 2018 the US Army had 58 Brigade Combat Teams, of which 16 were Armour against 9 Strykers (South, 2018)

5 Conclusion

Our application of the information gap approach to the problem of allocating resources for national security yields three broad conclusions, each of which is evaluated quantitatively. First, the higher the robustness to model misspecification policy makers demand, the higher the overall level of defense expenditures required. Second, the greater the demand for robustness, the greater should be the reliance on those defense measures better suited to prevent extremely high levels of damage. Similarly, the more robustness policy makers wish to achieve, the more they should eschew investment in systems that promise the best *expected* outcomes on the battlefield, but are also more vulnerable to catastrophical damage when they fail.

Beyond the normative recommendations for how policy makers can best allocate resources to protect their countries from aggression, in a world in which much of the relevant information is unreliable, the info-gap method also provides some positive insights into the reasons for the policies policy makers choose today or have chosen in the past. In a world with unreliable probabilistic information, we should not be surprised if policy makers lavish higher expenditure on defense than would be appropriate if the only goal were expected utility maximisation. Furthermore, we would expect policy makers to favor expenditures on weapons systems, and associated tactics and strategies, that are both most effective in preventing worst-case scenarios and are also better understood. These would suggest one possible rationale for military planners' reputation for conservatism, and indeed inertia, when confronted with new and untried technologies and doctrines.

A century before Frank Knight made the distinction between risk and uncertainty, Clausewitz noted the problems associated with the unreliability of information in military settings.

> Many intelligence reports in war are contradictory; even more are false, and most are uncertain .... One report tallies with another, confirms it, magnifies it, lends it color, till [the officer] has to make a quick decision which is soon recognized to be mistaken, just as the reports turn out to be lies, exaggerations, errors, and so on. In short, most intelligence is false, and the effect of fear is to multiply lies and inaccuracies .... The general unreliability of all information presents a special problem in war: all action takes place, so to speak, in a kind of twilight, which like fog or moonlight, often tends to make things seem grotesque and larger than they really are. (von Clausewitz, 1993, pp. 136, 161)

The information gap approach does nothing to ameliorate the unreliability of information, but it does offer policy makers a methodology to make decisions better suited for such environments.

Appendix A Derivation of the Robustness Function

We derive the robustness function for an horizon of uncertainty not exceeding unity, $\alpha \leq 1$. We make no assumptions about the utility function $u(c)$ other than that the marginal utility is positive: $u'(c) > 0$.

The main task is to find the pdf of the damage, $p(\psi|\chi)$ which, at horizon of uncertainty α, minimises the expected utility $R(\chi|p, P_{\mathrm{w}})$ defined in (1). Because the marginal utility is positive it is evident that $R(\chi|p, P_{\mathrm{w}})$ is minimised by that pdf in $\mathcal{F}(\alpha, \widetilde{p}, \widetilde{P}_{\mathrm{w}})$ which assigns as much weight as possible to large levels of damage and as little weight as possible to low levels of damage. For the fractional-error info-gap model in (10), one readily shows that $M(\alpha, \chi) = \min R(\chi|p, P_{\mathrm{w}})$ in (7) occurs with the damage densities $\hat{p}(\psi|\chi)$ given in (14) and (15), for $\alpha \in (0, 1]$ and $\alpha > 1$, respectively.

Consider first the case $\alpha \leq 1$. The utility $R(\chi|p, P_{\mathrm{w}})$ in (1), evaluated with the pdf in (14), is:

$$R(\chi|p, P_{\mathrm{w}}) = (\widetilde{r} - \delta_r\alpha - u(c))\, P_{\mathrm{w}} + u(c) \tag{24}$$

where $\widetilde{r}$ and δ_r are defined in (11) and (20) and $u(c)$ is the utility if an attack does not occur. The term $\widetilde{r} - \delta_r\alpha - u(c)$ is negative, since positive marginal utility implies that $\widetilde{r}_1 > \widetilde{r}_2$, and therefore $\delta_r > 0$, and $\widetilde{r} < u(c)$. Consequently the value of P_{w} in $\mathcal{F}(\alpha, \widetilde{p}, \widetilde{P}_{\mathrm{w}})$ which reduces the right-hand side of (24) to its minimum is $(1+\alpha)\widetilde{P}_{\mathrm{w}}$. Thus the minimum expected utility, up to horizon of uncertainty $\alpha < 1$, is (see (7)):

$$M(\alpha, \chi) = (\widetilde{r} - \delta_r\alpha - u(c))\,(1+\alpha)\widetilde{P}_{\mathrm{w}} + u(c) \tag{25}$$

Equating the right-hand side (25) to R_{c} yields a quadratic equation in α, the positive root of which is given by:

$$\alpha_0(R_c, \chi) = \begin{cases} \dfrac{(2\widetilde{r}_2 - u(c))\widetilde{P}_{\mathrm{w}} + \sqrt{\left(2\widetilde{r}_2 - u(c)\right)^2 \widetilde{P}_{\mathrm{w}}^2 + 4\delta_r \widetilde{P}_{\mathrm{w}}\left((\tilde{r} - u(c))\widetilde{P}_{\mathrm{w}} + u(c) - R_c\right)}}{2\delta_r \widetilde{P}_{\mathrm{w}}} & \text{if } R_{\mathrm{c}} \leq \widetilde{R}(\chi) \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

For instances where $\alpha \geq 1$:

$$M(\alpha, \chi) = ((1+\alpha)\widetilde{s}_2 - u(c))\,(1+\alpha)\widetilde{P}_{\mathrm{w}} + u(c) \tag{27}$$

Following the same procedure yields an *implicit* relationship for the positive root:

$$\alpha_1(R_c, \chi) = \begin{cases} \dfrac{\left(u(c) - 2\widetilde{s}_2\right)\widetilde{P}_{\mathrm{w}} + \sqrt{[u(c) - 2\widetilde{s}_2]^2 \widetilde{P}_{\mathrm{w}}^2 - 4\widetilde{s}_2 \widetilde{P}_{\mathrm{w}}\left[\widetilde{P}_{\mathrm{w}}\widetilde{s}_2 + (1 - \widetilde{P}_{\mathrm{w}})u(c) - R_c\right]}}{2\widetilde{s}_2 \widetilde{P}_{\mathrm{w}}} & \text{if } R_{\mathrm{c}} \leq \widetilde{R}(\chi) \\ 0 & \text{otherwise} \end{cases}, \tag{28}$$

where:

$$\widetilde{s}_2 \quad = \quad \int_{\psi_\alpha}^{1} u\left[(1-\psi)(1-\chi_1-\chi_2)\right]\widetilde{p}(\psi|\chi)\,\mathrm{d}\psi. \tag{29}$$

For each particular allocation χ there exists a value of R_c we denote as R_c^1 such that $\alpha_0(R_c^1,\chi)=\alpha_1(R_c^1,\chi)=1$, where the values of ψ_{m} and ψ_α coincide and and hence so do the values of $\widetilde{r}_2$ and $\widetilde{s}_2$. For values of $R_c \le R_c^1$ $[R_c > R_c^1]$, $\alpha_0(R_c,\chi) \ge \alpha_1(R_c,\chi)$ $[\alpha_0(R_c,\chi) < \alpha_1(R_c,\chi)]$ and hence $\widehat{\alpha}(R_c,\chi)=\alpha_0(R_c,\chi)$ $[\widehat{\alpha}(R_c,\chi)=\alpha_1(R_c,\chi)]$. Together this implies that $\widehat{\alpha}(R_c,\chi) = \max[\alpha_0(R_c,\chi),\alpha_1(R_c,\chi)]$. We note that the values of both $\alpha_0(R_c,\chi)$ and $\alpha_1(R_c,\chi)$ are declining in R_c. This can be directly seen in (26) for α_0. it is important to emphasise that unlike (26), (28) does not represent a closed-form solution for α_1 because unlike ψ_{m} in (14), the value of the quantile ψ_α, in (15) and (16) and hence $\widetilde{s}_2$ in (29) as well, are themselves non-linear functions of α. Therefore for $\widehat{\alpha} \ge 1$, the relationship between $\widehat{\alpha}(R_c,\chi)$ and R_c can only be explored numerically.

Appendix B Example Specifications of the Policymaker's Best Estimates of Model Components

This subsection presents the functional forms that we use in our subsequent illustration. These specifications were chosen to reflect the historical magnitudes described above, and to capture the intertwined channels through which defense allocations affect both the probability of suffering an attack and the losses such an attack would impose. The best estimate of the pdf of damage, conditional on being attacked, is assumed to be the Beta function:

$$\widetilde{p}(\psi|\chi) = \frac{\psi^{a(\chi)-1}(1-\psi)^{b(\chi)-1}\Gamma(a(\chi)+b(\chi))}{\Gamma(a(\chi))\Gamma(b(\chi))}, \tag{30}$$

where the Gamma function is given by $\Gamma(x)=\int_0^\infty t^{(x-1)}e^{-t}dt$ and:

$$a(\chi) \quad = \quad 1+\frac{\chi_1}{\chi_2}+\theta e^{\theta\chi_1}\ln(\chi_1+\chi_2)\,\chi_1, \tag{31}$$

$$b(\chi) \quad = \quad 2+\frac{\chi_1}{\chi_2}+e^{\chi_2}\ln(\chi_1+\chi_2)\,\chi_2. \tag{32}$$

The mean of this pdf is:

$$\mathrm{E}(\psi|\chi) = \frac{\Gamma(1+a(\chi))\Gamma(a(\chi)+b(\chi))}{\Gamma(a(\chi))\Gamma(1+a(\chi)+b(\chi))} \tag{33}$$

We set the value of $\theta = 3$ in our calculations.

The best estimate of the probability of suffering an attack is:

$$\widetilde{P}_{\mathrm{w}}(\chi) = 1-\left(\frac{\beta_{\frac{1}{2}}(a(\chi),b(\chi))\Gamma(a(\chi)+b(\chi))(\chi_1+\chi_2)}{\Gamma(a(\chi))\Gamma(b(\chi))}\right)^{\frac{1}{5}} \tag{34}$$

where $\beta_x(a,b) \equiv \int_0^x t^{a-1}(1-t)^{b-1}dt$ and $a(\chi)$ and $b(\chi)$ are defined in (31) and (32). The term $\chi_1+\chi_2$ represents the total military expenditure, and reflects its deterrent value. In our example there is an *positive* relationship between the expected damage (conditioned on being attacked) that corresponds to a particular allocation of defense spending χ and the probability of being attacked associated with that allocation. This can be seen in the term $\frac{\beta_{\frac{1}{2}}(a(\chi),b(\chi))\Gamma(a(\chi)+b(\chi))(\chi_1+\chi_2)}{\Gamma(a(\chi))\Gamma(b(\chi))}$ in (34) and represents the probability that damages will be small—less than half the maximal possible damage—conditioned on attack taking place. Hence, the higher this number is, the lower the likelihood that an adversary will be tempted to launch an attack. Finally, we adopt the constant-risk-aversion specification for the utility function $u(c) = c^{1-\gamma}/(1-\gamma)$, where $\gamma > 0$.[14]

References

T. K. Adams, *The Army After Next: The First Postindustrial Army*, Praeger Security International, Westport CT, 2006.

Y. Ben-Haim, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, Academic Press, London, 2nd edition, 2006.

Y. Ben-Haim, Info-gap Decision Theory For Engineering Design. Or: Why 'Good' is Preferable to 'Best', in E. Nikolaidis, D. Ghiocel and S. Singhal, eds., *Engineering Design Reliability Handbook*, CRC Press, Boca Raton, 2005.

S. Biddle, "Speed Kills? Reassessing the Role of Speed, Precision, and Situation Awareness in the Fall of Saddam," *The Journal of Strategic Studies*, 2007; 30: 3-46.

V. Bier, S. Oliveros and L. Samuelson, "Choosing what to protect: strategic defensive allocation against an unknown attacker," *Journal of Public Economic Theory*, 2007; 9: 563-587.

M. Boot, *The New American Way of War, Foreign Affairs*," *Foreign Affairs*, 2003; 82: 41-58.

D. L. Brito, "A dynamic model of an armaments race," *International Economic Review*, 1972; 13: 359-375.

R. J. C. Butow, *Tojo and the Coming of the War*, Stanford University Press, CA, 1961.

E. A. Cohen, "A Revolution in Warfare," *Foreign Affairs*, 1996; 75: 37-54.

A. J. Echevaria II, "Enduring Elements of Military Strategy," in P. M. Cronin ed., *The Impenetrable Fog of War*, Praeger Security International, London, 2008.

[14] In our calculations we set $\gamma = .98$, so utility is close to logarithmic.

T. Farrell, S. Rynning, and T. Terriff, *Transforming Military Power Since the Cold War: Britain, France, and the United States, 1991–2012*, Cambridge University Press, 2013.

A. Futter, "Conclusion: Reflecting on the RMA Concept," in A. Futter and J. Collins (eds.), *In Reassessing the Revolution in Military Affairs*, Palgrave Macmillan, London, 2015.

A. I. George, *Presidential Decisionmaking in Foreign Policy: The Effective Use of Information and Advice*, Westview Press, Boulder, CO., 1980.

M. I. Handel, *War, Strategy and Intelligence*, Frank Cass, London, 1989.

V. D. Hansen, *The Second World Wars*, Basic Books, New York, 2017.

D. Hume, *An Enquiry Concerning Human Understanding*, Dover Publications Inc., Mineola, NY, 2004.

T. L. Hughes, *The Fate of Facts in a World of Men*, Headline Series, No. 233, December 1976, Foreign Policy Association, Inc., New York.

M. D. Intriligator and D. L. Brito, "Proliferation and the Probability of Nuclear War," *Public Choice*, 1981; 37: 247-260.

F. Kagan, *Finding the target: The transformation of American military policy*, Encounter Books, New York, 2007.

J.E. Kaufmann and H.W. Kaufmann, *The Maginot Line: None Shall Pass*, Praeger Security International, Westport CT, 1997.

J.E. Kaufmann and H.W. Kaufmann, *Fortress France: The Maginot Line and French Defenses in World War II*, Praeger Security International, Westport CT, 2006.

Y. Niho and M. Takeuchi, Optimal Defense Spending, in T. Negishi, R. V. Ramachandran ed. and K. Mino *Economic Theory, Dynamics and Markets*, Springer Science, New York, 2001.

L. F. Richardson, *Arms and Insecurity*, Boxwood, Pittsburgh, 1960.

P. W. Singer, *Wired for war: The robotics revolution and conflict in the 21st century*, Penguin Books, New York, 2009.

T. South, The Army is converting two BCTs as it beefs up its fighting force for the next big war," *Army Times*, 20 September, 2018. Retrieved from: https://www.armytimes.com/news/your-army/2018/09/20/the-army-is-converting-two-bcts-as-it-beefs-up-its-fighting-force-for-the-next-big-war/

D. Talbot, Technology Failed in Iraq," *MIT Technology Review*, 2004; 107: 36-45.

W. J. Thies, "Alliances and Collective Goods," *Journal of Conflict Resolution*, 1987; 31: 298-332.

Third Infantry Division (Mechanized), *After Action Report, Operation Iraqi Freedom*, 2003. Retrieved from: https://www.globalsecurity.org/military/library/report/2003/3id-aar-jul03.pdf

F. van der Ploeg and A. J. de Zeeuw, "Perfect Equilibrium in a model of competetive arms accumulation," *International Economic Review*, 1990; 31: 131-146.

C. von Clausewitz, *On War*, trans. Michael Howard and Peter Paret, Everyman's Library, London, 1993.