Costly Peace: A New Rationalist Explanation for War*

Andrew J. Coe

September 21, 2011

Abstract

I argue that scholars of war have neglected an important rationalist explanation for war: sometimes peace is more costly than war. This explanation is important because it provides sounder accounts of some historical wars than other explanations. To demonstrate this, I identify three empirically common sources of costs in peace: arming, imposition, and predation. For each, I provide a simple model that demonstrates the conditions under which war can occur and show that these wars are not due to other rationalist explanations. I then offer analytic narratives of the Iraq War, the civil conflicts in Iraq after the Gulf War, and the American Revolution based on these models, and argue that these accounts fit the facts better than other rationalist accounts.

*I am grateful to Robert Powell for a series of conversations that inspired this line of thinking. Obviously, he should not be held responsible for its realization here. I also thank James Fearon, Jeffry Frieden, Robert Powell, Dustin Tingley, and the participants in various workshops at Harvard University for comments on earlier versions of this essay.

1 Introduction

Why do wars happen? Since wars are destructive, it is not obvious why they would be chosen over peaceful bargaining as a means of resolving contested interests between rational actors. The bargaining theory of war therefore focuses on explaining why peaceful bargaining might fail and thereby lead to costly violence. This increasingly influential and wide-ranging body of scholarship focuses on two main causes of war: shifts in the balance of power between two disputants, which can lead a side that fears decline to attack the other in order to forestall its rise; and asymmetric information about a side's willingness or ability to resolve an issue through war, which may motivate one side to risk war in order to call the other's possible bluff. These two rationalist explanations have been applied to particular wars, whole classes of wars, and even specific aspects of war and international relations more generally.[1] They are becoming pervasive in the study of conflict by both political scientists and economists, to the exclusion of many alternative explanations for war (Jackson and Morelli, 2009).

This essay is intended to convince you that a third rationalist explanation for war—costly peace—is of comparable importance to the other two, despite its relative neglect by scholars. The essence of this explanation is that the anticipated costs of peace may exceed those of war. Rational actors will then make demands of each other that cannot be mutually satisfied, because the total value of any feasible peace is less than that of war. Thus, bargaining will fail and war will occur, even in the absence of any shift in power or asymmetry in information.

This is not merely a theoretical possibility. There are several empirically common sources of costs in peace, and this essay will argue that there are historical wars that are more cogently explained by costly peace than by the other rationalist explanations. Thus, this explanation is important because it is relevant to understanding the origins of empirical wars.

This essay explores three sources of costly peace, termed arming, imposition, and predation. For each, a simple model is provided that exposes the way in which this source can make peace costly and lead to war. Each model is then used to construct an analytic narrative of an empirical war, and evidence is presented that these wars derive from the costs of peace introduced by each

[1] For recent reviews, see Jackson and Morelli (2009); Powell (2002); Reiter (2003).

source.[2]

First, actors may take expensive measures, most commonly arming, in order to defend their claims to disputed stakes. If this costs enough over the longer term, they may prefer to fight immediately in order to avoid paying these costs. It will be argued that the US-Iraq War that began in 2003 arose from the costs of maintaining the balance of power between the belligerents. More generally, wars of consolidation, secession, and succession may derive substantially from the costs of arming.

Second, peace may mean the imposition of penalties on, or removal of rewards from, two potential adversaries by outside actors. If these are severe enough, war may result as actors fight to obtain these external gains. It will be argued that the civil conflicts in Iraq that followed the Gulf War of 1990–91 arose from the sanctions imposed by the international community in the aftermath of the war. More generally, powerful countries often impose sanctions or offer rewards explicitly aimed at inciting regime change, violently if necessary.

Finally, peace sometimes entails the transfer of resources or the exaction of tribute from one side by another. This predation undermines incentives for productive activity, and so is costly. If the costs are high enough, war may occur as one side tries to lessen the other's predation and the other asserts control. It will be argued that the American War of Independence was fought because Britain had begun to prey on the thirteen colonies, and the colonists expected more. More generally, wars from ancient Rome through to 19th century America were fought to discourage or eliminate piracy and raiding.

To be clear, theorists of war have known of this third rationalist explanation for some time. Powell (2006) points out the possibility that the costs of arming might alone suffice to cause war. Powell (1993, 1999) analyze models of arming in which the costs of arming might cause war. McBride and Skaperdas (2007) uses this mechanism to explain why conflict happens in low-income countries, McBride and Skaperdas (2009) tests it in a laboratory experiment, and McBride, Milante and Skaperdas (2011) explores how good institutions might mitigate it. Fearon (2011) uses this mechanism to provide a theory of the democratic peace. Slantchev (2010) models a related

[2]For more on the method of analytic narratives, see Bates et al. (1998).

mechanism, whereby the need in peacetime to repay debt taken on to increase military strength can lead to war. Bas and Coe (2011) analyzes nuclear proliferation, showing that wars over the spread of nuclear weapons are most likely to arise from the costs of containment. And finally, Powell (2011) shows that the imposition of outside benefits for decisive outcomes can also cause war.[3]

What is missing from this body of research on costly peace is a *demonstration of its ability to explain real wars*. The primary contributions of this essay are to provide three models of war due to costly peace that are tailored to explaining why specific historical instances of war happened, and to show that these accounts fit the facts better than others. It also offers analyses of the three models that are designed to expose the underlying common logic of war due to costly peace, and identifies and analyzes predation, a source of costly peace that has not previously appeared in the literature. Finally, it offers suggestions intended to help researchers in applying this explanation to other wars.

The next section explains costly peace and its relationship to war, situating it within the bargaining theory of war and explaining some previously unnoticed implications of costly peace for the broader understanding of war. Section 3 presents a model of war due to the costs of arming that is tailored to the context of the interaction between the United States and Iraq after the first Gulf War. It then constructs an analytic narrative of the Iraq War based on the model, and evaluates the performance of this model against other explanations for the war. Section 4 does the same for imposition and the civil conflicts that took place in Iraq between the Gulf War and the Iraq War. Section 5 does the same for predation and the American War of Independence. Section 6 concludes with suggestions for how to determine if the costs of peace could have caused a particular war, and points out additional candidates for such wars.

[3]A line of work following Hirshleifer (1991) analyzes the distortionary effects of predation, but does not make the connection to war and the failure of peaceful bargaining explicit.

2 Costly Peace and War

It is helpful to begin by explicating the quintessential elements of the bargaining theory of war and its way of framing war as an "inefficiency puzzle." With these elements in hand, I will explain what costly peace is and how it can lead to war. I then examine its relation to the other rationalist explanations for war.

2.1 Cheap Peace and the Inefficiency Puzzle of War

The bargaining theory of war models a situation in which two rational actors have opposed interests over some stake, whatever it might be, and may attempt to revise its extant disposition.[4] They have just two means of doing so: they can bargain in an attempt to come to a revision that is implemented by mutual agreement, or they can fight to impose a revision unilaterally. Bargained revision (peace) is assumed to be free; violent revision (war) is assumed to be costly.

From this perspective, war is mysterious. Why would war ever occur, when the participants could simply peacefully implement the expected outcome of a war, thereby avoiding its costs and doing strictly better? When the actors are risk-neutral over the contested stake, as is usually assumed, any peacefully agreed outcome is efficient, while war is inefficient because it imposes costs. Thus, the "inefficiency puzzle" of war is why actors would choose an inefficient means of resolving a dispute when an efficient means is available. Even if actors are allowed to be risk-averse, then peace may not be efficient, but it is still always "cheaper" overall than war.

The crux of this framing is that peace is cheaper than war, not because of some deep empirical law, but *by construction.* The seminal models of the theory are designed so that fighting is the only thing that imposes costs on the players: there is nothing else they can do, and nothing else that can happen to them, that will destroy value. This probably explains why costly peace was initially missed as a rationalist explanation for war. With a very few exceptions, even the many extensions of the seminal models, dealing with domestic politics, intra-war bargaining, diplomacy, and many other aspects, nonetheless retain this property.[5]

[4]The exposition of the bargaining theory of war given here relies heavily on Fearon (1995).

[5]The exceptions will be discussed in subsequent sections.

This construction has been very useful in advancing the theory, but it is importantly flawed. As subsequent sections will demonstrate, there are several empirically important sources of costs in peace. Each of the models presented there makes use of a simple property of costly peace: when the costs of peace are higher than those of war, then war must occur. To see why this is true, we first need to talk about what costly peace is.

2.2 When Costly Peace Leads to War

Peace is costly if it involves one or both actors taking measures other than war that nonetheless destroy some of the value—whether wealth or any other desirable thing—that could otherwise be realized by the two. Arming is one such measure: it destroys value because resources that could be consumed and enjoyed are instead devoted solely to protecting one actor's share of the total value or taking more of the other's. Another is that one actor might not agree to a settlement that would bring in external rewards; by doing so, he imposes the loss of these rewards on both. And finally, if one actor produces less because the other will prey on the fruits of his labor, then there will be less goods and service to consume.

War might lead to a reduction in these various costs. For instance, if one actor is vanquished, the other will no longer have to arm against him to protect his share of the value. In the case of imposition, if the truculent actor is defeated, the other can freely implement a settlement that would reap the external rewards. And if the predator is eliminated, then there is no more discouragement to production and so no loss of consumption. Of course, war is itself costly in that it also destroys some of the value available to the players. So the costs of war can be thought of as including both the direct cost of fighting, and also any costs of peace that will remain after the war.

If the anticipated costs of peace are greater than the anticipated costs of war, then fighting would be expected to leave more value for the surviving actor(s) to enjoy than the two combined would take from peace. That is, war would increase the value available to the actors over what they could realize from peace. When this is true, there is no peaceful deal that both actors would prefer to war. Any deal that gave one actor as much value as he would expect to get from war would leave the other actor short of his own war value. Because of this, there is no way to divide

up the value of peace so that both actors are satisfied.

If, by contrast, the anticipated costs of peace are less than those of war, then fighting would be expected to leave less value to the actors than settlement. That is, war would decrease the value available to the actors relative to peace. When this is true, there are settlements that would divide up the value of peace so that both actors got at least their war values. This does not guarantee the actors would not fight, as costly peace is not the only rational cause of war, and if peace is costly, then it might exacerbate another cause of war. But it does mean that, when this condition holds, the costs of peace alone do not suffice to cause war.

Thus, costly peace is sufficient to cause war if and only if the costs of peace exceed the costs of war. This fact is true, but vacuous, for the seminal models of the bargaining theory of war. In those models, as in all the models to be presented here, the costs of war are always positive because war entails destruction. But with the typical assumption of risk-neutrality, the costs of peace are always zero by construction, because in these games there is nothing the players can do that destroys value other than fight. Even if risk-aversion is allowed, the costs of peace are always strictly less than those of war, because the latter then entails not only destruction but also costly uncertainty. Thus, the costs of war exceed those of peace, and war can never happen due to costly peace in these models.

It is enlightening to to restate this fact in terms of efficiency: costly peace leads to war because it means that peace is even more inefficient than war. From this perspective, the inefficiency puzzle of war also applies to costly peace.[6] If the actors could simply not take the measures that make peace inefficient, then peace would be efficient and thus yield more value than both costly peace and war. But then there is a way to divide the value of peace so that both actors would get more value from this division than from costly peace or war. If this is true, why would actors choose either to make peace costly or to fight?

The answers are just the same as those used in the bargaining theory of war: commitment problems (CP) and asymmetric information (AI). At root, the rationalist explanations for war offered by Fearon (1995) are actually rationalist causes of *inefficiency*, whether it comes in the

[6] Slantchev (2010) makes this observation as well.

form of war or of costly peace.[7] If one of these causes is present, then the occurrence of costly peace or war simply depends on the availability and costs of each.

One might then conclude that, theoretically speaking, there is nothing new here. If CP and AI cause costly peace, just as they cause war, then what have we learned? One implication of taking costly peace into account is that, while CP or AI might be necessary causes for war to occur, their presence alone does not suffice. If there are costly measures other than war that the actors could take in response to CP and AI, then the presence of these causes might lead only to costly peace. War will happen only if these measures are ineffective or too costly. The seminal bargaining models of war missed this because they rule out such measures by assumption.[8]

2.3 Implications of Costly Peace

Empirically, there are in fact often costly measures other than war that actors can employ in response to deal with both CP and AI. In some cases, these are sufficiently cheap and effective so that war can be avoided; in others, war is cheaper. An immediate implication is that, to fully explain an actual case of war, one needs not only to determine the ultimate source of inefficiency (whether CP or AI or both), but also why costly measures other than war were not employed instead. A second implication is that, to the extent that applications of the seminal models are used to make predictions, these will be biased toward predicting war, because the possibility of other costly measures to address the underlying CP and AI is ignored.

A third, and more unsettling, implication is that in some cases wars may actually improve the welfare of the belligerents relative to what would happen if they remained at peace. If, in the absence of war, the actors would take costly measures to address underlying issues of CP and AI, then peace will be costly. If it is more costly than war, then war would improve their social welfare, and so preventing war between these actors might actually make them worse off. This is radically different from the implications that follow from models of war where peace is always efficient, and

[7] Fearon (1995) offered not only commitment problems and asymmetric information as rationalist explanations for war, but also issue indivisibility. However, Powell (2006) shows that this last explanation in fact relies on a commitment problem.

[8] Of course, these models were designed to pose the inefficiency puzzle of war in the starkest possible terms, and for that purpose assuming away costly peace is a valuable simplification.

so bears some explanation.

In the absence of costly peace, there is necessarily more value to go around before a war than after. Thus, there is always at least one actor who is made worse off by war in expectation and so regrets the war the moment it starts—because of this, war without costly peace can never be a Pareto improvement. To illustrate with the three causes of war from Fearon (1995), in the case of shifting power, the regretful player is the one who expects to gain advantage in the future. With asymmetric information, the regretful player is the one who is surprised by an unexpectedly strong adversary. With issue indivisibility, there is always one player who would receive more than his war value in peace, and this player has reason to regret a war.

By contrast, because a war due to costly peace increases the total value available to the actors, at least one of them must be left strictly better off by war. It is even possible that, in expectation, *both* actors will do strictly better from war.

The crux of this difference is that wars in the absence of costly peace are entirely about one actor's willingness to pay to prevent a too-generous portion of the total value from being allocated to the other. Thus, regardless of who wins, these wars merely shift value between actors. The problem is that this shift comes at a cost: the total value is reduced because the war is costlier than peace. This is why at least one actor must be left worse off. Before the war, the expected allocation favors one player, who would like to prevent war but cannot because of budget, credibility, or divisibility constraints, and who thus regrets the war. In contrast, wars due to costly peace are not about altering the allocation of value, but about increasing the total value available. It is this increase in value that makes a Pareto improvement from war possible.

This property of wars due to costly peace calls for a fundamental re-examination of the policy advice that typically follows from bargaining models of war. It implies that there may be some wars—those due to costly peace—that are in the actors', and possibly the world's, best interest. Thus, unlike wars in the absence of costly peace, it may not always be true that wars due to costly peace should be prevented or stopped.

Developing this implication is beyond the scope of this paper, but I can give one example to substantiate the claim that a re-examination is needed. With wars in the absence of costly peace,

it is always true that increasing the costs of war by enough will cause the actors to forego war and thus leave them better off overall. (Increasing the costs of war by too little just means that war still happens but is more costly for the participants, making them worse off.) Thus, the application of sanctions, pressure for a ceasefire and mediation, and introduction of peace-keeping forces are all steps that, if they are strong enough to prevent or stop a war, should be taken. However, with wars due to costly peace, increasing the costs of war enough to take war out of equilibrium can actually leave the participants worse off if it does not also decrease the costs of peace enough. This implies that an ability to distinguish whether a particular ongoing war was primarily due to costly peace would be extremely valuable for policy-makers. I'll return to this problem in the final section of the essay, as it is also relevant for scholars looking to make use of costly peace in explaining historical wars.

2.4 A Simple Example of Costly Peace and War

The subsequent sections of this essay will substantiate the empirical relevance of these ideas. Three historical wars will be examined using models that explicitly incorporate costly measures that actors can take in response to CP and AI. I will argue that in all three cases, the available measures would have effectively eliminated the difficulties associated with CP and AI in these contexts. And I will endeavor to show that these measures were abandoned in favor of war because they were more expensive than war. However, because these models are tailored to represent actual historical contexts, they introduce many features that improve realism but obscure the underlying logic described above. So I will close this section by working through a highly simplified example, intended to make the ideas discussed in this section a little more concrete.

Consider the relationship between costly peace and two other well-known, though rarely-used, rationalist explanations for war: risk acceptance (RA) and issue indivisibility (II). Briefly, if one or both actors are sufficiently risk-acceptant over the disputed stake, then war will occur in equilibrium because it settles the dispute via a gamble, which the risk-acceptant actor(s) strictly prefer(s) to implementing the expected outcome peacefully. Issue indivisibility causes war if all of the

settlements of the disputed issue that both actors would prefer to war are actually infeasible.[9]

First observe that these two explanations are related. According to Fearon's (1995) definition, if a contested stake is completely indivisible, then the only possible settlements are those where one actor receives the entire stake. However, there is another conceptualization of issue indivisibility that is consistent with Fearon's story and also his examples: if a contested stake is completely indivisible, then any non-trivial division of it will give both actors zero value. Seen this way, issue indivisibility is just an extreme version of risk-acceptance, in which the actors like risk so much that anything other than an all-or-nothing gamble is worthless to them.

Now, Powell (2006) shows that wars due to RA and II are caused by a commitment problem. Suppose the actors were able to commit to determine the disposition of the contested stake via a weighted coin toss, where the weights correspond to the actors' probabilities of victory in war. Then both would strictly prefer to do so rather than fight, because they would get the same outcome in expectation, but without paying the costs of war. But of course they cannot commit—there is nothing to stop the player that loses the toss from refusing to honor it.

We can refine Powell's claim using the ideas discussed in this section. In the case of RA, the commitment problem is a necessary cause for war, but it is not sufficient to cause war on its own. If the level of risk acceptance is low enough, so that actors only get a small premium from gambling, they will not fight. But in the case of of II, the commitment problem is both necessary and sufficient for war. Why?

The difference is that, when actors are risk-acceptant but the issue is divisible, there is a costly measure other than war they can use to address the commitment problem: they can compromise

[9]Risk acceptance is usually disregarded as an explanation for war on the grounds that most of the actors to which the models are applied do not appear to be generally risk-acceptant, and current psychological theory does not offer a sufficiently precise characterization of the circumstances in which individuals will become risk-acceptant to be empirically prescriptive in the case of war. For instance, prospect theory states that individuals will be risk-acceptant when in the "domain of losses," but the definition of this domain relies on an individual's "reference point," a parameter that is difficult to specify *a priori*. Thus, in applying the theory to war and peace, reference points must be chosen in some *ad hoc* fashion, undermining the generalizability and utility of the resulting hypotheses. See [cite Renshon]. It is thus unenlightening empirically and uncompelling theoretically to simply assume that actors are risk-acceptant. The subsequent models presented in this paper will assume that actors are not risk-acceptant.

For its part, the empirical relevance of issue indivisibility has been questioned on the grounds that most disputed stakes are continuously divisible or at least amenable to alternating control, and any others are not so valuable that side payments cannot be made to compensate one actor for ceding possession of the stake and thereby avoid war (Fearon, 1995).

on a (sure) division of the stake. This is costly, because it entails losing the risk premium. When the smallest compromise that will appease either actor is costly enough, the costs of peace with this compromise will exceed the costs of war, and war will result. By contrast, when the issue is completely indivisible, compromise is impossible, and war results because there is no other way to address the commitment problem.

To see this, assume that two actors play a one-shot game in which A (the "proposer") goes to war or makes an offer q to B (the "receiver") on how to settle an issue represented by the unit interval $[0, 1]$, wherein A gets q and B gets $1 - q$, after which B either accepts the proposal or goes to war. If both choose peace, A's proposal is implemented and the game is over. If either chooses war, then with probability $p \in (0, 1)$ A receives the whole interval, and otherwise B does, and either way each player pays a cost $c > 0$. We will assume for convenience that the players' war payoffs are positive, so that $p - c, 1 - p - c > 0$. A's valuation of the settlement q is $u_A(q) = q^n$ and B's is $u_B(q) = (1 - q)^n$, for some $n \in \mathbb{N} \cup \infty$.

Notice that the players are risk-acceptant so long as $n > 1$, and increasingly so as n increases: settlements strictly between 0 and 1 are worth increasingly less than their "face value." This implies that $u_A(q) + u_B(q) < 1$ for all $q \in (0, 1)$, so that any intermediate settlement is costly relative to either player getting the whole interval. In the limit as n goes to infinity, the issue becomes completely indivisible: all settlements other than 0 or 1 give a payoff of zero to both players.

Suppose first that $n < \infty$. Then the players have a reasonable option of compromise over the issue: while costly, it does not eliminate the whole value of the issue. The best candidate for a peaceful compromise is the one that minimizes the costs of peace while doing *just* enough to appease one of the players, as this maximizes the value left to appease the other player. Suppose that B is the stronger player, so that $p \leq 1/2$. Then it is easily shown that the best candidate is $q^* = 1 - \sqrt[n]{1 - p - c}$. This gives B a payoff that is equal to his war value, and leaves more value for A than an alternative compromise in which A got exactly his war value would leave for B.

Would this compromise be preferred to war? The discussion above suggests the answer is no whenever the costs of peace exceed the costs of war, or $1 - (1 - p - c) - \left(1 - \sqrt[n]{1 - p - c}\right)^n > c + c$.

Re-arranging this gives $p - c > \left(1 - \sqrt[n]{1-p-c}\right)^n = u_A(q^*)$, which just says that A's war payoff exceeds his value from this compromise. Thus, war will happen exactly when the costs of peace exceed those of war.

The more risk-acceptant the players are (the higher is n), the more likely the costs of peace are to exceed those of war. If $n = \infty$, then war cannot possibly be avoided. Both players' war payoffs are positive, but peace guarantees that one player will receive a payoff of zero, so one will always strictly prefer war. There is simply nothing else that can be done—*any* compromise is exorbitantly costly.

Thus, when costly measures to address the inability to commit to a peaceful gamble are available and cheaper than war, peace prevails. When these measures are unavailable (as in Fearon's definition of issue indivisibility) or too expensive (as in the one worked out here), then peace is costlier than war and war occurs.

3 Arming and the Iraq War

The first, and perhaps most obvious, empirical source of costs in peace is *arming*: expensive measures taken by either actor to improve its prospects in war or undermine its opponent's. It is well-known to bargaining theorists that if the anticipated costs of future arming exceed those of war, then actors will choose war because it is cheaper than sustained arming (Powell, 1999, 2006; McBride and Skaperdas, 2007, 2009; McBride, Milante and Skaperdas, 2011; Slantchev, 2010; Bas and Coe, 2011; Fearon, 2011). But it is not known whether this explanation for war is empirically relevant: is there a historical war that is plausibly due to the costs of arming?

In this section, I will resolve this question by giving an analytic narrative of the 2003 US-Iraq War (henceforth, the "Iraq War"). The narrative is based on a model of war and arming (in this case, containment) that is tailored specifically to the interaction between the United States and Iraq prior to the war. To apply this model, a set of assumptions must be made that calibrate the model to the empirical circumstances that pertained during the decade of US containment of Iraq. From these assumptions, a single relationship is derived that governs whether war occurs: if the costs of containment come to be perceived as exceeding the costs of war, then there will be war. The

analysis of the model yields a new account of the Iraq War—based on costly peace—that is quite different from existing accounts and that more closely fits the empirical record of what happened. It thereby demonstrates that the costs of arming is an empirically important explanation for war.

Since I will argue that the model applies directly to an actual war, the empirical accuracy, or at least theoretical innocuousness, of its features must be carefully described and defended. After doing this, I will state the calibrating assumptions to be used, and defend their empirical validity. There follows a proposition that characterizes the equilibrium outcome under these assumptions, which is proved in the appendix. We will discuss the intuition for this result: why costs arise in peacetime, how war is expected to reduce them, and why the actors could not otherwise avoid these costs. And finally, I will give a narrative of the Iraq War based on this analysis and compare it to existing accounts.[10]

3.1 A Model of the US-Iraq Interaction after the Gulf War

To begin constructing the model, suppose there are two players, the US and Iraq. Given the totalitarian rule Saddam Hussein exerted over Iraq, modeling the country as a unitary actor in this way is not a bad approximation. In the US case, there were of course many different influential actors involved in decision-making over Iraq, but in the end most executive officials, majorities of both houses of Congress, and a large majority of the American public supported the strategy the Bush administration pursued, so that the US can be taken as, in practice, acting unitarily.

The two players have conflicting interests over a set of issues, represented by the unit interval $[0, 1]$, which they must somehow divide between them over time. These issues include influence over other states in the region, control over regional oil reserves, relations with regional terrorist organizations, the treatment of the people of Iraq, and perhaps others. Any settlement of these issues is a point in $[0, 1]$; assume for convenience that the US favors settlements closer to one,

[10]There are a great many non-rationalist (or at least, not formally rationalist) accounts of the war, in terms of the personal characteristics of key decision-makers, the confluence of key ideas among important elites, dysfunctional bureaucracies, concerns over re-election, special interests, and misperceptions (Flibbert, 2006). While these accounts elucidate the many issues over which the US and Iraq had conflicting interests, they do not crystallize why the two states could not find a peaceful settlement that both would prefer to war. By contrast, the few existing rationalist accounts of the war, to be cited and described below, focus on explaining the failure of peaceful negotiation. Because this is also the focus of my narrative, I will compare it only to the other rationalist accounts.

while Iraq favors settlements closer to zero. Further, assume for simplicity's sake that both players are risk-neutral over these issues, and that both discount payoffs over time at a constant rate $\delta \in (0,1)$.[11] Thus, the utility of a sequence of payoffs from settlements, with q_t the implemented settlement at time $t \in \mathbb{N}$, is $U_{US} = \sum_{i=0}^{\infty} \delta^i q_i$ for the US and $U_{IR} = \sum_{i=0}^{\infty} \delta^i (1 - q_i)$ for Iraq.

The first period of the game can be thought of as the period immediately after the end of the Gulf War of 1990–1991. It begins with the US choosing whether to engage in containment, an action to be discussed momentarily. The US then chooses either to go to war against Iraq or to make a peaceful offer of a settlement for that period. Iraq can then accept the US offer, in which case it is implemented for that round, or reject it and go to war.[12] If Iraq accepts the US offer, then it can also choose whether or not to try to acquire nuclear weapons. If a peaceful settlement is agreed, then in every subsequent round, Nature moves first and determines whether Iraq's effort in the previous period to acquire nuclear weapons is successful, and whether the US detects this effort. If it is successful, then Iraq is assumed to have nuclear weapons for that and all future peaceful periods. If it is not, then Iraq must try again to have a chance of getting nuclear weapons in the next period. After Nature's move, the choices described above repeat. War is treated as a game-ending costly lottery, to be described shortly.

I am assuming here that the acquisition of nuclear weapons is an inherently uncertain endeavor. Many technologically-sophisticated inputs are required to produce a nuclear weapon, and successful mastery of each of these inputs takes an uncertain amount of time. Additionally, both containment and the attempt to acquire nuclear weapons are represented as binary choices—the US cannot choose *how much* to contain Iraq, and Iraq cannot choose how hard to try to get nuclear weapons. Allowing for continuous choices here would complicate the analysis considerably, but would not change the conclusions presented below, because the problem that leads to war does not derive from the inability of the US or Iraq to choose just the right degree of weapons development or

[11] These assumptions are standard in the bargaining theory of war, but it is worth noting the effects of relaxing them. None of the qualitative results derived here will change if I allow the players to be risk-averse, or to discount future payoffs at differing rates, so long as the discount is always positive and grows over time. Allowing the players to be risk-acceptant generates the possibility that war occurs purely due to the desirability of gambles, but the costs of arming that are the focus here would still contribute to the incentives for war.

[12] Here, for expositional simplicity, the US is modeled as making take-it-or-leave-it offers to Iraq. The choice of bargaining protocol will not affect the results, as they do not depend on the allocation of the bargaining surplus, but only on whether a surplus exists.

containment. I also assume that the US cannot react instantly to Iraq's efforts to acquire nuclear weapons—if Iraq tries to get them, there is some probability it will be successful before the US can launch a war. Even if the US instantly and accurately observed Iraq's efforts, it would take some time to mobilize itself for war, during which Iraq might be successful.

Most importantly, I will assume that the probability that governs Iraq's development efforts depends only on whether the US elects to contain or not—it cannot, for example, rise over time as Iraq keeps trying and masters various inputs to nuclear weapons. Allowing for this to occur when the US does not contain would strengthen the conclusions, because it would increase the incentives for both containment and war, but it is important for the results that the probability be bounded above when the US does engage in containment. This restriction can be defended on empirical grounds: even strident advocates in the US for war believed that Iraq's efforts could be held in check so long as the US did what was necessary to uphold the sanctions and remained willing to compel Iraq's periodic opening to inspectors (Pollack, 2002), so that any progress it made could be detected and reversed. It can also be defended on theoretical grounds: Bas and Coe (2011) models this interaction and shows that the US can effectively cap the probability that Iraq's efforts are successful over time. Since Iraq has the ability to credibly reveal its progress at any given time by easing access for inspectors, and the US becomes willing to go to war if its estimate of the progress (and thus, the likelihood of a subsequent shift in the balance of power) gets too high, it is always in Iraq's interest to avoid war by opening up to inspectors and reassuring the US that the probability remains low.[13] The model presented here abstracts away from these aspects in order to focus on the costs of containment.

War is assumed to have only two possible outcomes: a complete US victory over Iraq and a (much less probable) complete Iraqi victory over the US. Allowing for a larger range of possible outcomes would not qualitative alter the results, so long as there remains a substantial probability of

[13]Some accounts of the war argue that Saddam's need to maintain ambiguity about his weapons programs, especially about his ability to deploy chemical and biological weapons, in order to deter foreign enemies such as Iran and domestic insurrection, impeded his ability to signal clearly to the US that he had abandoned these programs (Baliga and Sjöström, 2008; Lake, 2010). However, it seems clear that Saddam would prefer fighting Iran or an Iraqi opposition to fighting the US, given the overwhelming military advantages of the US. So, given a choice between certain war with the US in order to maintain ambiguity about his weapons programs, and possible revolt or war with Iran upon revealing the absence of such capabilities, he would choose the latter. As I will explain later, this is exactly what Saddam did once the US threat of war became credible in 2002.

a complete US victory that would be expected to eliminate the costs of containment. The expected outcome of a war started in a given period, as well as the probabilities that in the next period Iraq will obtain nuclear weapons and that its efforts will be detected, depend only on whether the US engaged in containment in the given period and whether Iraq had previously acquired nuclear weapons. First suppose that Iraq does not have nuclear weapons. If the US does not contain, then the probability of US victory in war is p, the costs of war for the players are $d_{US}, d_{IR} > 0$[14], the probability that Iraq will obtain nuclear weapons in the next period if it tries to get them is $\lambda > 0$, and the probability that the US will detect Iraq's efforts is $\sigma > 0$.[15]

Containment, as the US and its allies practiced it after the first Gulf War, was a strategy designed to ensure that Iraq was unable to threaten or coerce its neighbors, that Saddam was unable to inflict massacres on the Shi'i in the south or the Kurds in the north, and that Saddam's regime was unable to reconstitute and expand its once-formidable pursuit of weapons of mass destruction, including nuclear weapons.[16] It consisted of a comprehensive package of measures to achieve these ends. General economic sanctions were placed on Iraq to starve the regime of hard currency to rebuild its military, including restrictions on Iraq's ability to sell its oil, and US forces were stationed at bases near Iraq to ensure a rapid response to any new provocation. The net effect of these measures was to increase the chance that the US would be successful in a new war against Iraq to $p^c > p$ and to lower the cost of fighting such a war to $d^c_{US} < d_{US}$. They also lowered the costs of war for Iraq to $d^c_{IR} < d_{IR}$, since they ensured that the next war would be quick and entail minimal destruction of Iraqi resources in the form of rebuilt military forces. Additionally, the enforcement of no-fly zones in the south and north of Iraq to protect the minorities living there, including suppression of Iraq's air defenses, further increased the US advantage over Iraq and lessened the costs should war recur. Finally, intrusive monitoring of Iraq's military, including repeated inspections of any sites suspected of WMD-relevant activity, and limited strikes as necessary to convince Iraq's regime of the seriousness of US concerns at

[14]For mathematical convenience, all war costs are taken to be per-period costs, so that, e.g., the total cost of war for the US is $d_{US}/(1-\delta)$.

[15]For simplicity, I ignore the possibility of a "false positive": US detection of cheating when none is taking place. Allowing for a small probability of this occurring would not qualitatively change the results.

[16]I rely here on the discussion of the objectives, elements, and effects of containment in Pollack (2002), chapters 2 and 3, and the assessment of their effects in Cordesman (2002).

particular junctions and compel its compliance with inspections, lessened the probability that Iraq would obtain nuclear weapons in the near future to $\lambda^c < \lambda$ and increased the chance that the US would detect such efforts to $\sigma^c > \sigma$. Finally, these measures were themselves costly, so that in each period containment imposes costs of $c_{US}, c_{IR} > 0$ on the players.

Now suppose instead that, at some period after the start of the game, Iraq obtains nuclear weapons. Clearly, this would increase Iraq's chance of prevailing in a new war, and also increase the anticipated costs of such a war for both players. With a nuclear-armed Iraq, the US chance of victory would be $p^n < p < p^c$, and the costs would be $d^n_{US} > d_{US}$ and $d^n_{IR} > d_{IR}$. Note here that I am abstracting away from the possibility of containment after Iraq had obtained nuclear weapons. It is certainly possible that the US would continue some of the pre-nuclear measures, such as the stationing of forces nearby, and perhaps even add some, such as bolstering the missile defenses of its allies in the region, but others might be ended, as Iraq would be unlikely to tolerate limited strikes intended to protect its minorities or compel access for inspectors. Regardless, so long as Iraq would continue to want nuclear weapons even if it realized the US would try to contain it after it acquired them, relaxing this assumption and allowing for costly containment after Iraq got nuclear weapons would only strengthen the conclusions presented below.

Overall, it is assumed that Iraq prefers war when it is nuclear-armed to war when it is not and the US is not engaging in containment, and prefers both to war when the US is engaging in containment. That is, $1 - p^n - d^n_{IR} > 1 - p - d_{IR} > 1 - p^c - d^c_{IR}$. Analogously, the US is assumed to prefer war under containment to war without containment, and to prefer both to war when Iraq is nuclear-armed, or $p^c - d^c_{US} > p - d_{US} > p^n - d^n_{US}$. I will also assume that, if Iraq has not gotten nuclear weapons and the US intends to go to war, then it is worthwhile for the US to engage in containment prior to starting the war in order to gain the concomitant advantages: $\frac{p^c - d^c_{US}}{1-\delta} - c_{US} \geq \frac{p - d_{US}}{1-\delta}$.

Finally, all of the parameters of the game are assumed to be common knowledge. Before turning to the analysis of the game, it is worth remarking on three implicit features. First, I ignore the question of Iraq's *other* weapons of mass destruction (WMD): biological and chemical weapons. Whatever the uncertainties surrounding them, the shift in expectations about war outcomes due to

these weapons was small, because the US military was, especially by the 2000s, perfectly capable of winning a war against Iraq under chemical and/or biological attack (Cordesman, 2002; Pollack, 2002, pp. 33–37 and ch. 11). Second, there is no incomplete information about the two sides' preferences. Both sides had had plenty of time to learn each other's interests by 2002; there was no uncertainty about Saddam's desire to obtain nuclear weapons, or the US's willingness to bear substantial costs to minimize the chance of his success (Director of Central Intelligence, 2002).

Third, there is no terrorism in the model. Iraq's support for regional terrorist organizations might be regarded as increasing the costs of peace between the US and Iraq, which would strengthen the results presented below. Iraq might also transfer WMD or relevant expertise to terrorists, but despite the Bush administration's public assertions, the intelligence community placed low probability on this occurring, except perhaps if Iraq's survival was directly threatened (Director of Central Intelligence, 2002). Perhaps the most well-informed advocate of the war, Kenneth Pollack, supported this assessment and also regarded terrorism as the least of the threats posed by Iraq (2002, pp. 153–158, 178–180). Moreover, there were good reasons to believe Iraq would otherwise be deterred from doing so by the danger of US retaliation for any terrorist WMD attack (Mearsheimer and Walt, 2003; Pollack, 2002). We will return to these issues when we consider alternative accounts of the war.

3.2 Calibration and Analysis of the Model

To derive a precise prediction from the model, it is necessary to make four calibrating assumptions that narrow what can happen in equilibrium. Though they are stated verbally to ease understanding, each is equivalent to a certain relationship among the various parameters of the model. These assumptions imply that the equilibrium outcome of the game turns on a simple comparison between the costs of containment and the costs of war. After discussing what the assumptions mean and defending their empirical validity, I will state a proposition that characterizes equilibrium, and then discuss the intuition for this result. The proposition is proved in the appendix, and the translation of each assumption into a condition on the model parameters is performed therein.

Calibrating Assumptions:

1. *If the US chose to tolerate Iraq's nuclear weapons efforts and gave Iraq nothing prior to its acquisition of them, then Iraq would try to get them and would not start a war.*

2. *The US would be better off going to war than allowing Iraq to freely pursue nuclear weapons, even if it gave Iraq no concessions at all prior to its acquisition of nuclear weapons.*

3. *Absent containment, Iraq would covertly pursue nuclear weapons under any US-offered deal.*

4. *Containment would decrease the probability of Iraq getting nuclear weapons enough that the US would be able to hold Iraq to its war value.*

Intuitively, the first assumption means that Iraq would choose to wait until it acquired nuclear weapons to challenge the US, even if in the meantime the US conceded absolutely nothing to its interests, so long as the US did not engage in containment. This is crucial to the rest of the analysis: it ensures that Iraq anticipates substantial gains from trying to get the weapons. In the appendix it is shown that this assumption is equivalent to the following inequality:

$$1 - p - d_{IR} \leq \frac{\delta\lambda}{1-\delta}(p - p^n + d_{IR} - d^n_{IR}) \tag{1}$$

The left side of the inequality is how Iraq would expect to do in a war, absent containment. The right side is the total future benefit for Iraq of having nuclear weapons, relative to what it would get without the weapons if the US offered it just enough to avoid war, weighted by the probability that its efforts to get them are successful.

Empirically, this condition surely held. Judging from the earlier Gulf War, the US military was far superior to Iraq's even prior to containment, and the costs of war for Iraq were substantial, so that the left side would be close to zero. Both the US and Iraq believed that Iraq's acquisition of nuclear weapons would substantially improve its ability to extract concessions from the US.[17]

[17]For evidence that Saddam saw nuclear weapons as giving him more than just a guarantee of survival, see Pollack (2002, pp. 175–178). Some have argued that such "offensive" gains cannot be acquired with nuclear weapons, even for an aggressive leader like Saddam (Mearsheimer and Walt, 2003). However, it is widely believed within the US policy community that the spread of nuclear weapons, especially to potential adversaries, erodes US power and influence—one need only look at any edition of the National Security Strategy of the United States from the Clinton or Bush (Jr.) administrations to confirm this fact.

Moreover, Iraq was very close to mastering the technology by the time of the earlier Gulf War, so that the probability of success in developing them was high. Even after the war and the years of containment following it, the knowledge remained and the program could be reconstituted in the absence of containment (Duelfer, 2002*b*).

The second assumption means that the US would prefer war to just ignoring Iraq, even if that meant making no concessions to Iraq prior to its acquisition of weapons. It is equivalent to a complicated condition given in the appendix, but if the US is taken to be fairly patient ($\delta \approx 1$), as seems likely, then the condition simplifies to roughly $d^c_{US} + d^n_{IR} < p^c - p^n$. The left side of the inequality is the surplus from avoiding war that the US would gain by tolerating Iraq's nuclear program and conceding nothing to it, and the right side is the shift in the balance of power from Iraq getting nuclear weapons. I have already argued that containment greatly enhanced the already-large military superiority of the US over Iraq, and reduced the anticipated costs of war, so that p^c was close to 1 and d^c_{US} was small. Moreover, for Iraq to perceive substantial gains from nuclear weapons, as assumed and justified above, it must be that p^n is substantially less than p^c (the balance of power shifts in Iraq's favor) and d^n_{IR} is not too large (Iraq can afford to assert the advantages deriving from its nuclear weapons). Thus it is plausible that the inequality held true empirically.

Some observers argued before the war that toleration (and deterring Iraq once it got nuclear weapons) would be better for the US than war (Mearsheimer and Walt, 2003). Having seen the *ex post* costs of the war, many more would agree with this assessment now. However, the US had previously proven willing to run at least the risk of war to stop North Korea from developing nuclear weapons, though it ultimately failed (Mazarr, 1995, Chapter 8). Thus it seems reasonable to assume that, given a stark choice between tolerating Iraq's pursuit of nuclear weapons and fighting a war to stop it, the US would decide on the latter.

The second assumption ensures that the US will be willing to bear at least some cost to try to prevent Iraq's program from being successful. This cost could come in the form of war or containment, which would make the equilibrium inefficient because each destroys value, or a concession offered to Iraq in exchange for abandoning its nuclear program, which would be efficient because it simply transfers value from one player to another.

The third assumption means that the last option won't work: there is no way to avoid Iraq pursuing nuclear weapons without containment or war. If the US offered the most generous concession to Iraq that the US could tolerate in exchange for Iraq abandoning its program, and threatened Iraq with war if it was caught cheating, then Iraq would agree to the deal but still cheat on it. Anticipating this, the US would be unwilling to offer any such deal. From the appendix, this assumption is equivalent to:

$$(1-\lambda)\sigma\left[\frac{d_{US}^{c}+d_{IR}^{c}}{1-\delta}+c_{US}+c_{IR}\right]<\lambda\left[\frac{p^{c}-p^{n}-d_{US}^{c}-d_{IR}^{n}}{1-\delta}-c_{US}\right] \tag{2}$$

The left side is the surplus from avoiding containment and war that would be lost if Iraq was caught cheating on the deal, weighted by the probability that its covert effort was both unsuccessful and detected. The right side is the gain to Iraq from obtaining nuclear weapons, relative to the generous concession offered by the US to secure the deal, weighted by the probability that its covert effort succeeded. When the inequality holds, Iraq will discount the threat of US punishment, because it is unlikely to get caught before it gets the weapons, and even if it did, the punishment is not so bad.

Empirically, this condition very probably held. No serious analyst doubted that Iraq would reconstitute its nuclear program if containment lapsed, and throughout the era of containment there was little discussion of a deal of the kind considered here. Looking at the condition from the model, it is easy to see why. As already argued, the first and second assumptions imply that Iraq's chance of getting nuclear weapons absent containment (λ) was high, and that the bracketed term on the right side is positive. Containment's reduction of the costs of war (as described earlier) implies that the bracketed term on the left is not too large, but the most important factor there is σ, the chance that Iraq would get caught in the absence of containment. There was every reason to believe that this probability was quite small. Iraq had come close to acquiring nuclear weapons before the earlier Gulf War without the intelligence community or IAEA noticing. Even after years of containment, with its forced and intrusive monitoring, inspectors were unaware of major elements of Iraq's WMD programs, and only discovered them upon the chance defection of the director of

these programs, a member of Saddam's family (Duelfer, 2002*a*).[18]

Finally, the fourth assumption means that containment works well enough that Iraq has relatively little to gain from pursuing nuclear weapons under it, because this effort is so unlikely to succeed. Thus, to avoid Iraq starting a war, the US may have to make a (small) concession to Iraq's interests to make up for the fact that the advantages Iraq would get from acquiring nuclear weapons are unlikely to be realized any time soon. It also means that, under containment, the US is not interested in offering Iraq additional concessions in exchange for a promise to abandon its nuclear program. Any deal that was generous enough to win Iraq's compliance would be strictly preferred by Iraq because of the additional concession it entailed, but because containment works so well, it would not lessen the probability that Iraq would eventually get nuclear weapons enough to make it worthwhile for the US to offer this concession. In terms of the model, this assumption is equivalent to:

$$W_{IR}^{c} \geq \frac{-c_{IR} + \delta\lambda^{c} V_{IR}^{n}}{1 - \delta(1 - \lambda^{c})} \tag{3}$$

The left side is Iraq's expected value from war under containment; the right side is the value to Iraq of trying to get nuclear weapons under containment while being given nothing by the US. The lower Iraq's chance of success under containment (λ^{c}), the smaller the right side will be relative to the left.

The evidence available as of the early 2000s suggested that λ^{c} was very low. After all, starting from 1991, when it had been very close to getting the weapons, a decade had passed under containment and Iraq still did not have nuclear weapons. Moreover, many believed this was not for lack of trying—throughout the 1990s, the intelligence community had reporting from human sources that Iraq's program was continuing (Iraq Intelligence Commission, 2005, pp. 53–55). While this turned out to be false, it would nonetheless surely have contributed to a US perception that λ_{c} was quite low. Even in 2002, when Iraq had had four years without inspections to make progress, and the erosion of international sanctions had increased the resources available for pursuing nuclear

[18] As Pollack (2002, pp. 75–76) explains, this defection was not motivated by any foreign enticement, but rather by an improbable quarrel within Saddam's family that threatened the personal safety of the director. Thus, this highly unlikely revelation should not be considered as indicating good monitoring of Iraq's program.

weapons, the US intelligence community estimated that Iraq was still five to seven years away from getting a weapon (Director of Central Intelligence, 2002).

In reality, Iraq's ability to reconstitute its nuclear program decayed steadily after containment began, so that λ_c moved toward zero, in part because of Iraq's decision to temporarily abandon WMD in order to motivate the lifting of sanctions (Iraq Survey Group, 2004, Key Findings). But the bottom line is that the US had every reason to believe that containment had greatly reduced Iraq's ability to develop nuclear weapons, as indeed it did, and the fourth assumption thus held true.

Together, these assumptions imply that the equilibrium outcome must be either containment or war. The fourth assumption means that Iraq is indifferent between these; either way, it just gets its war value. Because of this, if war gives the higher total value for both players then both would prefer it to containment, and vice versa. Equivalently, if war imposes less costs on the players than containment, then it will occur. More formally, we have:

Proposition 1. *Under the calibrating assumptions, war is the unique equilibrium outcome if and only if the costs of containment exceed the costs of war. If the costs of containment are less than the costs of war, the unique equilibrium outcome has the US engaging in containment and Iraq pursuing nuclear weapons.*

Remarkably, the proposition implies that the equilibrium outcome is costly (that is, inefficient), whether it is peaceful or not.[19] The US and Iraq would do better overall if they neither fought a war nor engaged in containment. It is important to understand what causes this inefficiency, and why war occurs under the stated condition. There are three features of the model that are necessary for inefficiency to occur. First, Iraq cannot commit not to take advantage of nuclear weapons to obtain a more favorable settlement from the US once it has them. Second, the US cannot perfectly and costlessly police Iraq's pursuit of nuclear weapons. And third, absent costly containment by the US, Iraq cannot commit not to pursue the weapons.

The first commitment problem is exactly the same as the problem that leads to war in the

[19]The statement of the proposition ignores the knife-edge case where the costs are equal, in which there is an equilibrium with war and an equilibrium with containment.

standard models of shifting power (Fearon, 1995; Powell, 2006). If Iraq could commit not to take advantage of newly-acquired nuclear weapons, and did so, then the US would not care whether Iraq pursued nuclear weapons or not, and there would be no reason to engage in containment or fight a war. And if Iraq could commit, then it would certainly do so in equilibrium, because the US would be able and willing to entice it to do so by offering it part of the surplus gained from avoiding the costs of containment or war. Play would thus be efficient.

Since Iraq cannot commit not to take advantage of nuclear weapons if it gets them, the US does care whether Iraq tries to develop them. If the US were able to perfectly and costlessly monitor Iraq's efforts, and able to react instantly to the initiation of such efforts, then it could threaten to react to Iraq's efforts with war or containment. Faced with this threat, Iraq would be deterred from pursuing nuclear weapons, because the instant it did so, the US would detect it and take action to stop it—there would be no chance for Iraq's efforts to bear fruit, and thus nothing to gain from them. So, with perfect monitoring and instantaneous reaction, the commitment problem described above could be avoided, because Iraq would not pursue nuclear weapons and thus no shift in power could occur. Then, neither war nor containment would be needed and play would be efficient. Unfortunately, this possibility is ruled out in the model because, without containment, it is too easy for Iraq to conceal its efforts and takes too long for the US to react to them once detected.

Even imperfect monitoring and delayed reaction on the part of the US would not matter were it not for the second commitment problem. If Iraq could commit not to pursue nuclear weapons, and did so, then US monitoring of and reaction to Iraq's efforts would be irrelevant, as would be Iraq's inability to commit to not taking advantage of the weapons once it had them, and there would again be no reason to engage in containment or go to war. Once again, Iraq would make this commitment in equilibrium, because the US would entice it do so with a share of the surplus from avoiding the costs of containment or war, and play would be efficient.

Since Iraq cannot commit not to use nuclear weapons to coerce the US once it has them, the US cannot perfectly and costlessly police Iraq's nuclear efforts, and Iraq cannot commit not to pursue the weapons, the US is willing to employ costly measures to stop it from trying and

being successful. This ensures that the equilibrium will be inefficient: the only issue is whether it is peaceful (containment) or not (war). Containment is costly, but it lessens the probability that Iraq's nuclear program will be successful to the point that the US no longer expects it to happen, and so renders moot both the US inability to perfectly police Iraq's nuclear efforts and Iraq's commitment problems. In effect, the availability of containment ensures that neither shifting power nor asymmetric information are sufficient to cause war in equilibrium.

Similarly, war is also costly, and also renders irrelevant the inability to police and the commitment problems. If the US wins, then it can govern Iraq itself or install leaders with preferences similar to its own; either way, the dispute is ended and there is no more need to worry about Iraq's pursuit of nuclear weapons. If Iraq wins, then the US is (presumably) rendered unable to interfere further in Iraq's affairs, so that Iraq gets its way on the disputed issues and freely develops nuclear weapons if it so desires.[20]

The only reason that war would be chosen over containment is that the latter would be so costly that going to war would give the players a higher overall value. From the US perspective, going to war and "solving" the Iraq problem once and for all might be preferred to the indefinite continuation of expensive containment. Thus, in the model, war happens due to costly peace. To summarize, equilibrium is costly because of the commitment problems and inability to perfectly police, but it is only violent if the best candidate for peace—containment—is too costly.

3.3 An Analytic Narrative of the Iraq War

The reason the Iraq War happened is that the costs of containment grew rapidly through the 1990s and into the early 2000s, and were set to escalate further, while the anticipated costs of war steadily declined through the same period. Eventually, the costs of containment came to be perceived as exceeding those of war, and so, as Proposition 1 predicts, the US went to war. To substantiate this explanation, I will first describe the components of the costs of containment, and how each changed over time. After doing the same for the anticipated costs of war, I will review the available

[20]Note that given the overwhelming superiority of US forces relative to Iraq's, the probability of Iraq's victory might be approximately zero, so that assumptions about what would happen if Iraq won are essentially without loss of generality. It was widely believed that the US would surely win the war, but for a detailed assessment, see Cordesman (2002).

quantitative estimates of the costs of war and of containment. I will then consider why the war happened in 2003, in particular, and how the war was sold domestically. Finally, I will discuss some other explanations for the Iraq War and the evidence for them.

3.3.1 The Increasing Costs of Containing Iraq

The many costs of the US strategy of containment are described in Table 1. The table is divided into six components of the total cost: military, economic, humanitarian, diplomatic, political, and security costs. For each component, the principal costs are listed. To the left of each component is a symbol that indicates whether it generally rose (+), fell (−), or was approximately constant (0) over the course of containment from 1991 to 2003. The table is an original but obvious compilation of the costs of containment described across many sources, but relies heavily on Pollack (2002).[21] However, the trends specified for each component of costs deserve further explanation.

First, the military component of the costs of containment had increased over time, though only modestly. The principal rise was in the exposure of US forces stationed in the region to terrorist attack: the growth of Al-Qaeda raised the perceived risk, especially after the Khobar Towers attack of 1996 and the bombing of the *USS Cole* in 1998.

Second, the economic costs rose substantially over time. The sanctions imposed on Iraq led to a long, deep depression that reduced its economy to a shadow of its pre-Gulf War strength (Nordhaus, 2002, p. 58). The costs of lost access to Iraq's economy, especially to its oil, increased dramatically from 1999 as the price of oil rose quickly. Moreover, the costs increased for the US because its firms were made to respect the sanctions, while their competitors in Russia, France, and China were given leave to violate the sanctions in exchange for bribes from Iraq under the Oil-for-Food Program.

Third, the root cause of much of the increase in the costs of containment lay in the steadily increasing humanitarian costs of the sanctions. As Iraq's economy contracted, much of the remain-

[21] This might seem suspect, given that Pollack was a strong advocate for the war. However, the actual measures of which the US strategy of containment toward Iraq was composed, as discussed in the model setup, were a matter of public record and uncontroversial. Given a list of these measures, most of the specific costs listed in the table would be familiar to, for instance, anyone who regularly read the international section of the New York Times during the years of containment.

Trend	Component	Specific Costs
+	Military	committing forces to region; operations to enforce sanctions; strikes to compel compliance with inspections; enforcement of no-fly-zones; sporadic mobilizations to respond to Iraqi mobilization; exposure of troops to terrorist attack
+	Economic	trade and investment opportunities lost due to sanctions, including access to oil; loss of market share to defectors from sanctions; contraction of Iraqi economy
+	Humanitarian	impoverishment of Iraqi people; repression of Iraqi people; deterioration of Iraqi civil society; casualties inflicted by strikes
+	Diplomatic	recurring negotiations to maintain sanctions; coordinating responses to lack of compliance
+	Political	unpopularity and perceived illegitimacy of sanctions; appearing to persecute Muslims; corruption of UN bureaucracy by oil-for-food program; deterioration of relations with defectors from sanctions; domestic unpopularity of basing troops in hostile areas; inducement of terrorism by US military presence in holy lands
+	Security	of Iraq's support for terrorism; of foregone opportunity for catalyzing regional liberalization; of potential for terrorist use of Iraqi WMD; of induced vulnerability of Iraq to Iran

Table 1: The Costs of Containment of Iraq and Their Trends

ing surplus was appropriated by the regime to pay for the military and for the kickbacks necessary to undermine the sanctions through Oil-for-Food. As a result, the Iraqi people suffered increasing deprivation throughout the years of containment. This was universally viewed as undeserved and was widely attributed to the sanctions.

Fourth, the diplomatic costs of containment increased substantially over time as international support for containment declined. Especially as many countries began to respond to concerns over the impoverishment of Iraq caused by sanctions, and as some countries began to cheat on the sanctions, the diplomatic effort and political capital that had to be spent on the part of the US to police and preserve the sanctions and generate support for firm responses to Iraq's provocations rose.

Fifth, and in part as a result of rising humanitarian concerns, the political costs of containment rose dramatically over time. The sanctions, and especially US efforts to preserve them, became increasingly unpopular abroad, and fueled perceptions that the US was persecuting Muslims, es-

pecially after the Second Intifada and 9/11. Relations between the US and the defectors from sanctions suffered, especially in the cases of Russia, France, and China. The UN bureaucracy was increasingly undermined and corrupted by the dirty dealings of the Oil-for-Food program. Moreover, basing troops in areas where they were subject to terrorist attack became increasingly unpopular with the US public.

Sixth, and finally, the negative security externalities caused by containment also grew. Iraq increased its support to Palestinian terrorism as a means of bolstering its public image with the region's Muslims. In part because of this, the continued survival of Saddam's regime was viewed by some as a major impediment to Middle East peace. As the region's most domestically unpopular regime, its survival entailed foregoing the best available opportunity to catalyze democratization and liberalization and, some thought, a truly durable peace in the region. Last, after 9/11, prior beliefs about the risk of extreme terrorist attacks were revised upward, and the possibility that whatever WMD Iraq had might somehow fall into the hands of terrorists, whether inadvertently or intentionally, received greater weight.

Overall, the costs of containment rose quite substantially over the period 1991–2003.

3.3.2 The Decreasing Anticipated Costs of War with Iraq

Now consider the anticipated costs of war with Iraq, and how these changed over time. These are described in Table 2. The table is divided into the same six components as were used for organizing the costs of containment. For each component, the principal costs are listed; the general trend in each component from 1991 to 2003 appears on the left. The table condenses the highly detailed compilations of the actual and anticipated costs of war collected in Bilmes and Stiglitz (2006); Davis, Murphy and Topel (2006); Nordhaus (2002); Wallsten and Kosec (2005). As with the costs of containment, some explanation of the posited trends in the components of the anticipated costs of war is in order.

First, over the years of containment, the anticipated military costs of war had declined very substantially. Sanctions had drastically constrained the funds available to support Iraq's military, and its strength decayed rapidly. By contrast, the US military became much more capable over

Trend	Component	Specific Costs
−	Military	reduced availability of forces elsewhere; operations to support invasion and occupation; casualties and equipment destroyed in fighting; paid benefits for veterans; more difficult recruitment and retention; use of WMD
−	Economic	sharp rise in oil price; concomitant macroeconomic effects; destruction of civilian infrastructure in fighting; reconstruction of Iraq; increased risk imposed by war
−	Humanitarian	civilian suffering caused by war; potential for internal conflict after war
0	Diplomatic	marshalling international support for war and assistance with reconstruction; negotiating new US-Iraq relationship and status of forces in Iraq
+	Political	increased influence of Iran over Iraq; alienation of supporters of Iraq's regime; potential domestic unrest or punishment of incumbents if war goes badly; perception of attack on Muslims
−	Security	risk of expansion of war to Israel; potential new training ground for terrorists; externalities of post-invasion civil conflict, possibly including partition of Iraq, for surrounding countries

Table 2: The Anticipated Costs of War with Iraq and Their Trends

the 1990s—the development of high-precision bombing, dramatic improvements in inter-service coordination, and the introduction of modern information technology all radically increased the US military's ability to win wars. Moreover, the development of operational art featuring greater use of special operations forces and much-reduced manpower also promised to greatly reduce the cost of war, as it had in Afghanistan. Finally, the US military's ability to fight effectively under chemical and biological attack also improved substantially (Cordesman, 2002).

Second, the economic costs of war also fell. The same sanctions that strangled Iraq's military also devastated its economy. Oil production during the sanctions period, even after the initiation of Oil-for-Food, was only around half of production before the Gulf War. War would be less damaging to Iraq's economy because there was less infrastructure to blow up, and less room for economic activity to drop due to the disruption of war (Nordhaus, 2002).

Third, the humanitarian costs of war declined for the same reason as the economic costs. As the Iraqi people become more impoverished due to the sanctions, the room for further deprivation due to war was reduced. Additionally, the increasing military superiority of the US implied a shorter, more

surgical invasion which would minimize civilian casualties. Finally, there is no obvious rationale for believing that the potential for conflict within Iraq after the war changed substantially over time.

Fourth, the diplomatic costs did not change much over the years of containment. Those states that bore a larger share of the costs of containment—the UK and some of Iraq's neighbors—became more sympathetic to the case for war and restoring Iraq to working order after a war. However, the states that were making gains by defecting from sanctions became increasingly accommodating toward Iraq and opposed war (Pollack, 2002, Chapter 6 and pp. 352–365).

Fifth, there were competing trends within the political costs of war. First, the growing superiority of US forces meant less chance of international and domestic backlash from a long, destructive invasion. But, after 9/11 and the war in Afghanistan, the risk that a war with Iraq would be seen as further attacks on Muslim society, and thereby bolster support for international terrorism, increased. Also, as Iran recovered from the long, terrible war with Iraq of 1980–88, its ability to exert influence in Iraq, especially if Saddam's regime fell, grew. And as the economic links between the defectors from sanctions and Iraq grew stronger, the political costs of spurning these states' desire for the peaceful continuation of these links also rose. Overall, it seems plausible to assume that the political costs grew somewhat over time, but especially after 9/11.

Sixth, and finally, there were also competing trends within the security component. A better US military and weaker Iraqi military, especially with reduced missile capabilities, meant less risk of a war expanding to include Israel, or other neighboring countries. It also implied that the US would be more able to contain the externalities of any post-war civil conflict within Iraq. However, a war in Iraq, especially if it were followed by civil conflict, might provide a potential training ground for international terrorists, and this risk grew with the rise of Al-Qaeda. Still, as many observers have pointed out, this is potentially a net benefit, if it occupies terrorists in a location distant from the US and its allies and provides opportunities for their killing or capture by US forces.

Thus, most of the components of the anticipated costs of war with Iraq declined considerably over the period 1991–2003. It seems plausible to assume that, even at the end of this period, the combined fall of the military, economic, humanitarian, and security costs of war overwhelmed the rise in political costs, especially given the Bush (Jr.) administration's demonstrated insensitivity

to international political costs.

3.3.3 Comparing the Costs of Containment and of War

Immediately after the Gulf War, the Bush (Sr.) administration regarded war as more expensive than containment. That the costs of containment increased and those of war decreased over the subsequent years of containment does not imply that they ever crossed, so that containment became more expensive than war. So, here I will review the available estimates of the *absolute* costs of containment and of war, to see how they compare.

There are several potential pitfalls in this exercise. First, if the goal is to understand the US decision to go to war, then estimates based on information that became available *after* the war began are essentially irrelevant. Whether the war and subsequent occupation were competently managed, and whether this competence or lack thereof, or the occurrence of sectarian conflict in Iraq after the war, *should* have been predicted beforehand, are important questions, but they are distinct from the question of what *was* predicted. All that matters for this purpose are estimates of the costs of war and containment based on information available *before* the war.

Second, and relatedly, it should be clear from the preceding tables that many of the components, and indeed some of the most important components, of the costs of war and containment are exceedingly hard to measure. Even those which are naturally quantitative (e.g., economic costs) cannot be estimated with any precision. Thus, any attempt to total up the costs of containment and war is fraught with "imponderables" and "uncertainties" (Krueger, 2006).

Presumably for this reason, there is, to my knowledge, no publicly available estimate of the costs of war and containment that addresses all of the components identified in the preceding discussion. The most comprehensive estimates, on which I focus here, study only the military and economic costs in detail, and provide only back-of-the-envelope estimates of the humanitarian and security costs. The diplomatic and political costs are simply intangible—though no less real.

The restriction to reasonably comprehensive estimates based on *ex ante* information leaves just two studies. First, Nordhaus (2002) combines existing estimates from government sources and other economists with original calculations, based on information available as of late 2002, to estimate

the cost of war. This study has been praised for the soundness of its predictions and its frank treatment of uncertainty by the most comprehensive study of the war's actual (*ex post*) costs, and also by a post-war review of the pre-war estimates (Bilmes and Stiglitz, 2006; Krueger, 2007). It includes military and economic costs and, indirectly, some of the items listed as security costs in Table 2.

Second, Davis, Murphy and Topel (2006) (DMT) compiles a wide range of pre-war estimates of the costs of war and containment with original calculations, all based solely on information available before the war, to estimate the costs of a range of scenarios for both containment and war. It is the only study to estimate costs for both war and containment, and its figures include military, economic, humanitarian, and security costs.

The bottom line estimates are as follows. According to Nordhaus, the costs of war are predicted to range from approximately $100 billion to $2 trillion. According to DMT, the costs of war would be predicted to range from $100 to $870 billion, and the costs of containment would be from $300 to $700 billion. Moreover, the latter study also estimates that a war would improve the economic well-being of most Iraqis and their survival chances relative to containment.

There are several points to bear in mind when considering these estimates. First, though by focusing on the information available just before the war, they take account of the decline in the costs of war over the years of containment, they do not account for the trend in the costs of containment. That is, DMT assumes that the costs of containment in the future will be as they were in the past, averaged over those years, though their estimated range does include some low probabiity, unprecedented events such as the internal overthrow of Saddam. Second, both studies ignore the components of costs that are the most difficult to measure, and yet still produce ranges that vary by an order of magnitude. Moreover, both freely admit that these other components might well be of comparable magnitude to those they do estimate, and hence would increase the uncertainty substantially.

In a review of these estimates and others, Krueger (2006) takes a nihilistic view, observing that "all costs and benefits can be contested as wildly inaccurate—in either direction" and dismissing the exercise as "little more than educated guessing by other means." But whether these studies

should be taken seriously or not as estimates of the *true* costs of war and containment doesn't matter for our purposes. The fact that the costs of each option are so wildly uncertain could instead be taken as representative of an underlying, diffuse distribution of beliefs among the expert community about the costs of war and containment.

Interpreted this way, something *can* be learned from these estimates: there was a substantial overlap between the distribution of expert beliefs about the costs of containment and that of beliefs about the costs of war, even if the potential difference in either direction was great. Under these conditions, the beliefs of those in power about which option is better matter most. If these beliefs are strongly held, it is rational for those in power to use what means are available to convince the public to support their favored option, even if this entails shouting down experts with differing views or exaggerating aspects of their case.[22]

To separate the issue of why the war happened from how the case was made for it, it is helpful to examine why the war happened when it did, which I turn to next.

3.3.4 Why did the War Happen in 2003?

Every administration since the Gulf War eventually concluded that Iraq's regime had to be changed.[23] So, why is it that this was not done until 2003? An easy answer is that 9/11 happened, making the US public receptive to policy-makers' case for a war with Iraq, and war occurred as soon thereafter as the US was militarily prepared to invade. There are good reasons to believe that this was a necessary condition for the war to occur, but I will argue here that it was not sufficient. Two other conditions also had to be met, and the models developed in this section and the next expose the importance of these conditions. First, a third, cheaper option—inciting internal regime change—had to have been exhausted. Second, a large enough portion of the US policy community had to perceive that the costs of containment exceeded the costs of war. I'll elaborate on each of these two conditions, before re-considering the role of 9/11 in light of them.[24]

Containment was not originally designed to be a long-term strategy. At the end of the Gulf

[22]To be clear, by "rational" I do not mean to imply "moral" or "responsible." This essay is not intended to address moral concerns with the preparation, conduct, or consequences of the Iraq War.

[23]This fact is documented in Section 4.

[24]The historical material here is drawn from Chapters 2, 3, and 7 of Pollack (2002).

War, the US expected that Saddam's weakened grip on power would lead to his ouster by Iraq's generals and replacement with someone the US could tolerate. When this didn't immediately happen, the US implemented a policy of containment, which was intended to minimize the threat Iraq presented while the US waited for a coup. As part of this strategy, the US continued most of the unprecedentedly harsh sanctions from during the Gulf War and instituted a program of covert action, both intended to encourage the Iraqi opposition to depose Saddam. Unfortunately, Saddam proved rather talented at escaping a litany of coup and assassination attempts and suppressing several popular revolts. By 1996, all of the CIA-supported Iraqi opposition groups had been either broken or ejected from Iraq, and by 1999, further attempts were regarded as highly unlikely to be successful. In Section 4, I will provide a model that explains why, despite its failure, this strategy was reasonable, and document how views on this third option evolved. But suffice it to say that, by 2002, waiting for someone to depose Saddam was seen as a recipe for the indefinite continuation of containment.

While containment was highly effective at keeping Saddam in check, it came with all of the costs described above. Most importantly, it enabled Saddam to paint the sanctions, and their principal supporter the US, as the principal cause of the destitution of the Iraqi people. When, in response to international concern about the humanitarian externalities of containment, the Oil-for-Food program was implemented, Iraq thereby gained an indirect source of hard currency, in the form of a fixed but large quantity of oil contracts that could be allocated to buyers at its discretion. It used these contracts to separate France, China, and Russia from the US and UK, giving them privileged access to the Iraqi economy and threatening to end this access if they did not support its campaign to end containment.

Once this strategy of driving a wedge between members of the Security Council began to work, it became increasingly difficult for the US to enforce containment, as Iraq's provocations were now met with divided opinion and prevarication on the Council rather than firm consensus. With minor violations going unpunished, the floodgates were opened and many of Iraq's neighbors now began to violate the UN resolutions governing containment with abandon. The trend was clear by 1998, and in late 1999 the US consented to expanding Oil-for-Food to encompass a much broader array

of goods that Iraq could now import, essentially recognizing what was already taking place illicitly.

The Bush (Jr.) administration's first major foreign policy review upon entering office centered on Iraq. It decided to pursue the revitalization of containment with "smart sanctions." This policy offered to end the remaining economic sanctions on Iraq in exchange for renewed agreement to enforce restrictions on what goods Iraq was allowed to import, and offered incentives to the various defectors from containment that were intended to restore international consensus on the matter. Two attempts to move this initiative through the Security Council failed, because the incentives the Bush administration offered were not nearly enough to make up for the highly lucrative illicit trade in which many of Iraq's neighbors and supporters were already engaged.[25]

By 2002, it was clear that to continue effective containment, the US would have to take on *all* of the associated costs. Iraq's neighbors would have to be compensated by the US for the value of their illicit trade with Iraq. The US would have to unilaterally compel Iraq's compliance with the elements of containment. Its relations with France, China, Russia, and many of Iraq's neighbors, including US allies, would become more combative. Of course, this would only increase the general unpopularity of containment, as the US would no longer have international backing for its actions. Overall, the military, economic, diplomatic, and political costs of containment would rise dramatically.

Importantly, it is simply not true that it was impossible to restore containment, as some advocates for the war argued. Of course, containment would never be perfect: the US could never be *certain* that Iraq's WMD programs had been dismantled, any more than it had been for the decade of containment up to 2002. And it would cost much more, costs which, in the absence of a good prospect of inciting a successful coup or revolt against Saddam, would potentially have to be borne for a very long time. But it was still feasible. At the very least, the US possessed the military and economic resources necessary to contain Iraq unilaterally. It's just that this course of action would have been exorbitantly costly.

These two conditions—the exploding costs of preserving effective containment and the collapse

[25] A form of smart sanctions was eventually adopted in May 2002, but it excluded the anti-smuggling provisions, which were essential to the original design. Speculatively, if it is true that by that point the Bush administration had decided on war, then there was little point in continuing to fight for a better set of sanctions.

of the possibilities for inciting internal regime change—were perceived to hold as early as 1999 in the Clinton administration. Why didn't the US invade Iraq then? And if other challenges, such as the Kosovo War, the new round of Middle East peace negotiations, and the Monica Lewinsky scandal diverted the Clinton administration's attention through the end of its term, why didn't the Bush (Jr.) administration take up war when it entered office?

Certainly, many believed the new administration would do just that. It included, in substantial numbers and at very high levels, officials who had publicly and vehemently advocated a war to change Iraq's regime prior to entering government. And as mentioned, Iraq was the very first issue its foreign policy team considered. So why didn't it invade in 2001? A related question, for which the proposed answer is the same, is: why, in 2003, didn't the Bush administration just continue to kick the can down the road? After all, the US had been putting off serious action on Iraq since at least 1996, when the CIA's client groups in Iraq were defeated by the regime and the institution of Oil-for-Food began the erosion of containment.

The answer is 9/11, or rather the willingness it instilled in the American public to consider drastic action to improve US security. Since at least the years of US nuclear monopoly, when the United States considered war to prevent the Soviet Union from obtaining nuclear weapons, US policy-makers have perceived an unwillingness in the American public to countenance aggressive, unprovoked war to improve US security (Quester, 2000; Silverstone, 2007). According to Pollack, the Clinton administration's consideration of a US invasion of Iraq was always stymied by the belief that the public simply would not support such a war, at least not without grave provocation by Iraq. And even the earliest reported date at which the Bush administration had set a policy of invasion was still months after 9/11.

The Bush administration recognized the window of opportunity that the shock of 9/11 created. In order to gain the public support for doing what many of its members had long regarded as necessary with respect to Iraq, it needed only to develop a case for the war that would connect the two in the public's mind. There followed the allegations about the possible "nexus" wherein Iraq might transfer WMD to a terrorist organization for use against the US or its allies, the reported politicization of intelligence estimates, the silencing of official voices pointing out the likely high

cost of a war and occupation, and the claims that restoring containment was impossible. And, in time, the Iraq War.

Still, even if 9/11 was necessary for the Iraq War to occur, it does not fully explain it. Why did so many members of the Bush administration (and also the former Clinton administration, and Congress, and the independent policy community) see a war as desirable? It was not a matter of the new appreciation of the dangers of international terrorism, because these people had been advocating war well before 9/11. And why was war, rather than revitalized containment (for which the US surely could expect some generosity from the international community in the wake of 9/11) and a renewed commitment to inciting revolt, the chosen solution?

This essay offers answers to these questions. Internal regime chance was no longer regarded as a realistic option, and indefinite containment had grown more expensive than war. 9/11 simply provided the opportunity to overcome the public acceptance of the (eroding) status quo and aversion to the costs and risks of war. I will conclude this section by discussing the implications of this theory for other rationalist accounts of the Iraq War.

3.3.5 Implications for Other Rationalist Accounts of the War

Every other rationalist explanation for the Iraq War of which I am aware has at its core the same drivers of inefficiency as the model I have presented here.[26] First, there is asymmetric information about Iraq's pursuit of WMD (or its possession of WMD, or its intentions with respect to WMD). Second, there are commitment problems in that Iraq cannot commit to not developing WMD or to not using them once acquired. I have argued in this essay that, in general, such explanations for war are not sufficient unless analysts also demonstrate that measures other than war were unavailable, ineffective, or too costly.

This concern may seem superfluous here: after all, I have also argued that in the case of the Iraq War, these other measures (i.e., containment) were indeed too costly. However, ignoring the availability of these measures leads to interpretations of the Iraq War that do not do justice to the facts. Iraq's commitment problems and asymmetric information about its pursuit of WMD

[26] I know of only three: Baliga and Sjöström (2008); Debs and Monteiro (2010); Lake (2010).

were present from the beginning of containment, and so cannot explain why the war did not happen in the 1990s. One could use the argument of the previous section, that 9/11 created a window of opportunity for war, to justify the occurrence of war a decade later, but there were many more propitious windows earlier on. Iraq was caught red-handed with evidence of hidden WMD—including nuclear weapons infrastructure—more than once in the early 1990s. In the same period, the US repeatedly assembled broad international support for strikes against Iraq to enforce its compliance with containment. Why, when it had incontrovertible evidence of Iraq's perfidy and broad international support for aggressive action, didn't the US invade Iraq then?

One answer for why the war did not occur until after 9/11 is the so-called "one percent doctrine," promulgated by Vice President Cheney. This held that even a miniscule probability that a terrorist organization might obtain a weapon of mass destruction from Iraq and then use it against the US or its allies should be treated as though it were a certainty, because the consequences would be so extreme. There are a number of practical problems with this argument, including the lack of any reason to believe Saddam would risk such a thing, with which many members of the Bush administration would have been familiar. But there is also a deeper one: why is war necessary to eliminate this possibility?

More generally, it is not clear how asymmetric information about Iraq's pursuit or possession of WMD could lead to war. The development and stockpiling of biological, chemical, or nuclear weapons is not like resolve or the precise balance of military capability, parameters bargaining theorists often take to be asymmetrically known. It might well be impossible for one side to credibly convey its strength or resolve without fighting, because these things are inherently difficult to observe. But WMD and the physical capital and scientists that produce them are physical objects. Given the opportunity to look, it is always possible to see if they are present in a particular place. And it is possible for Iraq to allow the US to look freely for them, anywhere and anytime it likes, for as long as it likes, with as many inspectors and as much equipment as it likes.

It was thus in no sense impossible for the US to become approximately certain that Iraq had no weapons of mass destruction and was not actively pursuing them. In fact, it did just that in the aftermath of the war. The only thing preventing the US from achieving this certainty was Iraq's

refusal to comply with the inspections. However, if there's one thing on which the historical record is clear, it is that Saddam would do anything to preserve his regime. If he became convinced that unless total cooperation with inspections was forthcoming, the US would invade, then he would cooperate. Faced with the choice of a war he was almost certain to lose, and fully revealing and losing his WMD but having a chance to survive, he would surely take the latter.

Of course, Saddam never fully cooperated with the inspectors. He didn't need to: he only needed to give the US sufficient confidence in the rudimentary state of his programs to avoid an invasion. And, as we now know, even though he actually had virtually nothing to hide, there were other values for him in maintaining some ambiguity. Some have proposed this as an explanation for Saddam's refusal to do what was necessary to convince the US that he had abandoned his programs. That is, it is argued that Saddam was willing to run a risk of war in order to maintain the WMD ambiguity necessary to deter Iran and his internal enemies. But risk of war is not certainty of war. Faced with the choice between not cooperating and certainly losing a war with the US, and cooperating and possibly having to fight a much weaker enemy such as Iran or Iraqi rebels, Saddam's preference is still clear.

Thus, asymmetric information about the possession or pursuit of WMD could not cause war, because it could and would be dispelled at Iraq's behest whenever it felt the US threat of war was sufficiently credible. Asymmetric information about preferences is different, since these are not directly observable, but it is hard to believe that after more than a decade of constant bargaining and limited fighting, either the US or Iraq had any substantial uncertainty about the other's preferences. Hardly anyone in the US doubted that Iraq wanted to get WMD, or that it would accelerate its efforts if containment ended. The evidence from after the war suggests this belief was correct (Iraq Survey Group, 2004).

Instead, the US was perfectly capable of, if not verifiably ending Iraq's WMD programs, then at least assuring itself that they were like to be very small and, in the case of nuclear weapons, unlikely to succeed any time soon. The problem was that this assurance came at a cost: containment was not cheap. By 2002, after four years of rapidly eroding sanctions and without inspections, the CIA was uncertain about precisely what Iraq was up to. But the US rectified the situation by sending

a costly signal that it was serious about invading when it commenced a large buildup of military forces in the region in preparation for the war. Naturally, Saddam reacted by letting the inspectors back in, and Iraq cooperated more fully than it had in years.

The Bush administration was charged with ignoring the positive results of the inspections, with making unreasonable demands of Iraq, and with pressuring the intelligence community to produce more threatening estimates. These behaviors are all mysterious if the root of the problem was present uncertainty about Iraq's WMD programs. If the problem is lack of information, why wouldn't the results on inspections and the unvarnished estimates of the CIA be ameliorative? The answer is that both were actually irrelevant to the decision to go to war. In fact the problem was not asymmetric information, it was that the administration had already decided to stop paying the costs of containment and to go to war instead. Even if the inspectors had reported with confidence that Iraq had verifiably disarmed, and the intelligence had confirmed this, the true rationale for war would be unchallenged.

That said, every piece of evidence of Iraq's continued clandestine WMD programs and support for terrorism, no matter how circumstantial or suspect, would help to build support for the Iraq War among the public and the international community. But this evidence was solely useful for the sales pitch—and this explains why the administration would ignore the unfavorable (to its case for war) inspection reports and try to exaggerate the CIA's estimates.

The bottom line is that the facts are not consistent with accounts based solely on asymmetric information and shifting power (due to the future acquisition of nuclear weapons). Under containment, the asymmetry in information was quite limited, and the expected shift in power quite small, because the suspected small stock of chemical and/or biological weapons would not seriously shift the balance of power, and nuclear weapons were still a long ways off. Instead, the war occurred because maintaining this status quo had become more expensive than war.

The evidence presented here is by no means conclusive. But it does strongly suggest that an account of the Iraq War based on costly peace is more compelling than the existing accounts based on asymmetric information and shifting power. Thus, the costs of arming are demonstrably relevant to understanding the origins of at least one historical war.

4 Imposition and Civil Conflict in Post-Gulf War Iraq

Before the US went to war against Iraq in 2003, it tried an alternative method of avoiding the costs of containment: inciting a revolt to change Iraq's regime from within. Its approach was to impose sanctions on Iraq that were implicitly, and later explicitly, intended to encourage regime change. It is clear enough intuitively why the US might want to incite civil conflict to overthrow Saddam, but it is not immediately obvious why imposing sanctions would accomplish this. Why can't Saddam and the (potential) rebels come to a compromise that avoids the destruction of civil war or a coup? The answer, as I will demonstrate in this section, is that sanctions make internal peace between Saddam and the rebels costly; and if this peace is costlier than revolt, then a revolt will occur.

More abstractly, the second empirically common source of costs in peace is *imposition*: penalties imposed or rewards offered by outside actors that depend on the outcome of an interaction between two "inside" actors. These outside actors may be unwilling or unable to fight to affect this outcome, but they can still influence the choices made by the inside actors by making demands of them and offering incentives to meet these demands. If meeting these demands favors one inside actor but disfavors the other, the latter may refuse to do so, and peace will then entail the loss of the incentives offered by the outside actor. If these incentives are large enough, peace will be very expensive, and war will result as the disfavored side tries to hold its ground and the favored one tries to capture the outside incentives.

I am aware of only one model of imposition in the existing literature. Powell (2011) models a negotiation between two actors within a country over the balance of power between them. State consolidation, defined as the achievement of a near-monopoly of power, results in outside rewards that increase the value of the game. Consolidation can be achieved peacefully, but takes time and so delays the receipt of the rewards, making peace costly. If the rewards are large enough, then peaceful consolidation is more costly than war (i.e., violent consolidation), and war occurs. Empirically, the rewards are taken to be investment by foreign firms that is contingent on a sufficiently secure state. However, these outside actors are not modeled, and it is unclear why there is no deal they could work out with the inside actors that would enable investment to be made and war to be avoided, even in the absence of consolidation. Moreover, no empirical example of a war due to imposition

is presented, so it is unclear whether this explanation is actually relevant to explaining historical conflict.

Here, I will present a model of war due to the costs of imposition that is tailored to the situation *within* Iraq between the two wars with the United States. For tractability, this model ignores many of the aspects of the US-Iraq interaction considered in the previous section, but it endogenizes the design of sanctions on Iraq and the interaction between Saddam Hussein and an organized opposition (the inside actors) over how to respond to the sanctions. This approach has two advantages relative to that taken in Powell (2011). First, it illuminates not only why sanctions might cause civil conflict, but also why the US (the outside actor) might choose to incite this conflict, and thus why there is no acceptable deal among the outside and inside actors that would avoid it. Second, the model provides an explanation for the civil conflicts that took place in Iraq between the two wars with the US and for the role the US played in these conflicts. It thus shows that war due to the costs of imposition is empirically relevant.

After describing the setup of the model, I will state and provide intuition for the two results that characterize equilibrium. The section closes with a discussion of the aforementioned civil conflicts in Iraq.

4.1 A Model of Imposition and Civil Conflict

Suppose that Iraq is composed of two "internal" actors, labeled S for Saddam's regime and R for the opposition (rebels), who must bargain over control of the polity. The actors have opposed interests over, for example, whether Iraq is democratic or not, the degree of state control of the economy, minority rights and autonomy, the country's alignment with foreign nations, whether to pursue WMD, and perhaps other issues. These issues are represented by the unit interval $[0, 1]$; assume that Saddam favors outcomes closer to 1, the opposition those closer to 0, and that both are risk-neutral over outcomes in the interval, so that a settlement $q \in [0, 1]$ yields payoffs $u_S = q$, $u_R = 1 - q$ for the respective players.

The US also has an interest in these issues. Suppose the US is risk-neutral and favors outcomes closer to 0, just like Iraq's opposition, so that a settlement q gives the US a payoff of $u_{US} = 1 - q$.

Note that the outcome 1 does not represent the ideal set of policies for Saddam's regime—rather, it is the set of policies the regime would implement in the absence of any internal opposition, but still taking account of the need to appease the US. By contrast, 0 represents the policies the opposition would implement in Saddam's absence, and I assume that these are also the policies most favored by the US. Relaxing this assumption and allowing for there to be small differences between US and opposition preferences would not qualitatively alter the results presented here. Moreover, it seems reasonable given that the US took care to support only elements of the opposition that favored its interests (Pollack, 2002, Chapter 3). This support is not modeled explicitly, but to the extent that it is costly to the US and raises the rebels' chance of victory, incorporating it would strengthen the results.

The game described here is assumed to take place with the US already engaging in containment, having decided that this was preferable to a war with Iraq.[27] The game begins with the US choosing a demand to make and a sanction to impose if the demand is not met. The choice specifies the severity of the sanction $\alpha \geq 0$, and a target outcome $q_\alpha \in [0, 1]$. That is, if the internal actors do not implement a settlement in $[0, q_\alpha]$, the US will impose sanctions that cost Iraq α. Notice that this is equivalent, from the point of view of the internal actors, to a reward of size α in exchange for implementing an outcome no larger than q_α. Such a reward might include official development assistance, private direct investment, a security guarantee, access to foreign markets, diplomatic rehabilitation, and so on; the sanctions could entail the denial of any of these.

Imposing sanctions is costly for the US. The cost depends only on the severity of the sanction, and is given by the function $s(\alpha)$. It is assumed that $s(\alpha) > 0$, $s'(\alpha) > 0$, and $s''(\alpha) > 0$ for $\alpha > 0$, and that $s(0) = 0$, $s'(0) = 0$, and $s'(\alpha) \geq 1$ for high enough α. These ensure that: imposing non-zero sanctions is costly; the cost to the US of sanctions rises in their severity; the US implements the most cost-effective sanctions first; the first (tiny) measure of sanctions is effectively costless; and the marginal cost to the US of more severe sanctions eventually exceeds the marginal increase in their severity.[28] These costs are described in more detail in the previous section, but to remind,

[27] It would also suffice for the results if the US instead preferred war to containment, but also preferred inciting a revolt to war.

[28] The formulation given here says nothing about the possibility that there is a limit on the severity of the sanctions the US is capable of imposing, regardless of cost. In principle, the most the US could possibly do is to completely

they include economic costs (e.g., reduced international access to Iraq's economy), humanitarian costs (e.g., the impoverishment of Iraq's people), and diplomatic/political costs (e.g., the costs of maintaining international support for the sanctions, or of losing it and being regarded as acting illegitimately).

Since imposing sanctions is costly for the US, its threat to do so if its demand goes unmet may not be credible. Because the internal actors will not suffer the costs of sanctions if they wait until sanctions are actually imposed but then immediately agree to the US demand, they have no incentive to agree to it beforehand. Thus, if the US wants to exert influence over their choice of policies, it must actually impose the sanctions and pay the cost of doing so up front, before the internal actors make their decisions. So, I assume that upon making its demand and threat of sanctions, the US immediately implements those sanctions and pays the cost of doing so.

Once the US has specified its demand and sanctions, Saddam can either start an internal war or make an offer of a settlement of the contested policies to the opposition, which the rebels can either accept, in which case it is implemented, or reject, in which case there is war. War is a costly lottery that Saddam wins with probability p, the opposition with probability $1-p$, and which costs the internal actors $d_S, d_R > 0$ respectively. The winner is assumed able to implement his choice of policies within $[0, 1]$, while the loser gets nothing. It is important to note here that the structure of the game rules out the possibility that Saddam's regime might start a war with the US in order to stop the sanctions. This is consistent with the fourth calibrating assumption given in the previous section, which ensures that Iraq's value under containment is equal to its war value even given that the US is imposing quite severe sanctions.

So long as Saddam's regime remains in power, the US pays a cost c_{US}, which is the total cost of all the components of containment described in the previous section, except the costs of the sanctions. If the rebels take over, the US no longer needs to pay the cost of containment since the rebels share its interests, but a war within Iraq imposes a cost of $d_{US} > 0$ on the US, taken to be the humanitarian and diplomatic/political costs to the US of the suffering and destruction of a war

close Iraq to outside commerce of any kind. However, this possibility can be incorporated by simply assuming that the costs of sanctions become exorbitant above some upper bound, so that the formulation used here is completely general.

it had induced.

All parameters of the game are assumed to be common knowledge, and the game is solved for subgame-perfect equilibria.

4.2 Analysis of the Model

The analysis of the model is designed to answer two questions. First, under what conditions is the US *able* to use the imposition of sanctions to incite civil conflict within Iraq? Second, when would the US actually choose to do so? I will state a lemma and a proposition that answer these questions and discuss the intuition behind these answers; proofs are in the appendix.

Lemma 1. *If meeting the US demand is worse for Saddam than fighting an internal war, and the cost of sanctions for Iraq, weighted by the chance of ending them in war, exceeds the costs of war for Iraq, then war will occur. If meeting the US demand is better for Saddam than fighting, or the weighted cost of sanctions is less than the cost of war, there will be internal peace.*

The lemma answers the first question: the US would *always* be able to incite a revolt. It need only make a stringent enough demand and impose severe enough sanctions. If the demand is stringent enough, then Saddam would never voluntarily choose to meet it: he would spurn the demand even if it meant suffering a revolt. This means that if Saddam is left in power, the US demand will not be met and Iraq will have to endure the sanctions. Thus, the sanctions make peace between Saddam and the rebels costly. War is also costly, but it offers the possibility of ending the sanctions, should the rebels win, and so reduces the expected cost of peace. If the sanctions are severe enough, so that the costs of peace are higher than those of war, then the value left by the sanctions to divide up between Saddam and the rebels will not suffice to appease both, and war will occur.

Peace is costly because of a commitment problem identified by Powell (2006). In theory, Saddam and the rebels could eliminate the costs of peace relative to war by agreeing to set policies according to a coin flip weighted to reflect the internal balance of power. If Saddam won the toss, he would set his ideal policies and Iraq would endure the sanctions. If the rebels won the toss, they would set their ideal policies and the sanctions would end, just as if they had successfully revolted. The

expected costs from sanctions would then be the same in peace as in war, and thus there would be no need for war. Both Saddam and the rebels would commit to this coin flip if they could. But they cannot so commit: if the coin flip came out against Saddam, he could simply demand another toss, and the same is true for the rebels.

Proposition 2. *The US will incite a war within Iraq if and only if the cost of containment is high enough. Incitement becomes more tempting as the costs of war decrease, the chance of rebel success increases, and sanctions get cheaper.*

The proposition answers the second question: the US would choose to incite a war within Iraq when the cost of containment is too high. This is the same reason it would fight a war against Iraq in model of the previous section. There, the US was willing to pay the costs of a war with Iraq whenever these were less than the costs of containment. Here, the US is willing to pay the costs of *inciting* a war *within* Iraq—the cost of imposing severe enough sanctions, and the costs for the US of an Iraqi civil war—when these are sufficiently low relative to the cost of containment. When the costs of a civil war are lower, it is cheaper for the US to incite one, and when the rebels' chance of success is higher, an incited war is more likely to free the US of the cost of containment: both these increase the incentives to incite a war.

With the analysis in hand, we can turn to explaining the historical instances of rebellion that Iraq suffered after the Gulf War.

4.3 Sanctions and Revolt in Iraq

In what follows, I offer an analytic narrative of the civil conflicts within Iraq following the (first) Gulf War, guided by the model presented in this section. This model clarifies the origins of several, though not all, of these conflicts, and explains the US role in these conflicts and how it evolved over time. It also substantiates the claim made in Section 3, that the absence of a viable third option involving internal overthrow of Saddam partly explains why the Iraq War happened in 2003 rather than earlier. I rely heavily here on the more conventional historical narrative of the Gulf War and period of containment given in chapters 2 and 3 of Pollack (2002).[29]

[29]In this subsection, I will cite this book only with page numbers for the sake of brevity.

To begin, consider the problem faced by the United States at what became the end of the Gulf War. The Bush (Sr.) administration believed that removing Saddam and occupying Iraq would be extremely costly. Moreover, these costs might be avoided if someone else removed Saddam instead. The widely-held perception that the Gulf War defeat had seriously weakened Saddam's hold on power led the administration to expect that Iraq's generals would overthrow Saddam and replace him with someone more amenable to the US (Bush and Scowcroft, 1999, p. 488). If this happened, the US could then negotiate a new and more favorable relationship with Iraq. So, the administration actively encouraged it, with speeches broadcast and leaflets dropped into Iraq that encouraged the Iraqi military and people to rise up against Saddam (p. 48) and bring containment to a close.

Unfortunately, the post-war revolt against Saddam was launched not by the generals, but by the Shi'ites of southern Iraq, followed immediately by the Kurds of northern Iraq. The Bush administration feared that the success of this revolt would lead to a breakup of Iraq, the same fear that contributed to its decision to halt the Gulf War invasion short of Baghdad (Bush and Scowcroft, 1999, pp. 488–489). It also worried that the revolt would cause the Sunni elites at the center of power in Iraq to re-unite behind Saddam, as indeed it did, and thus reduce the chances of a military coup (pp. 48–49).

In terms of the model, we can think of the immediate post-war revolt as involving rebels whose preferences were perceived to differ substantially from those of the US. Though this was not modeled explicitly, it is equivalent to assuming that, should the rebels succeed, the policies they would then implement were distant from those preferred by the US. So long as the US thought that another set of rebels with interests more amenable to its own (e.g. a military conspiracy that would staunchly favor Iraq's unity) would be more likely to succeed, the success of the current revolt would make the US worse off. This explains why the US refrained from supporting it, even to the point of allowing Saddam the use of helicopter gunships to crush it. Pollack explains the lack of US support for this revolt in much the same way (p. 49).

When it became clear that Iraq's military had rallied behind Saddam to crush the revolt, the Bush administration proceeded to implement containment, with all its attendant costs. For all the

reasons described in Section 3, containment would protect US and allied interests from the threats Saddam posed while he remained in power. Moreover, and most importantly, it was expected to do so at a lower cost than full-on war to remove Saddam, the same reason the policy was reaffirmed later on by the incoming Clinton administration (pp. 65–66).

Importantly, the Bush administration's calculation of the cost of containment was predicated on the assumption that containment would not be needed for long (pp. 47–49, 52–53, 55). Senior officials still believed that it was only a short matter of time before a coup was attempted, and that it would be likely to succeed (p. 53, 55). The higher the probability of a successful coup, the lower the expected costs of containment, which need only be endured if the coup fails.

Why did the administration believe that another attempt at regime change was imminent? After all, Saddam had united the military and successfully put down the Shi'ite/Kurdish revolt, and was undoubtedly working to reduce the chances of such an attempt. The model explains why: containment included very severe sanctions, along with a stringent set of demands Iraq had to meet in order to end them. Obviously, Saddam would never voluntarily meet those demands, which included making restitution to countries harmed by the war (e.g., returning billions in stolen property to Kuwait and paying reparations) and abandoning Iraq's WMD and missile programs and support for terrorism. Thus, the sanctions gave any more amenable elites within Iraq enormous incentives to remove Saddam by force. In this way, sanctions make peace within Iraq costlier than revolt, and as Lemma 1 shows, revolt should follow.

The Bush administration well understood this, and did not even bother to uphold the pretense that the sanctions were about changing Iraq's policies, rather than its regime (p. 58). Moreover, the US was not about to sit back and hope to get lucky. As early as May 1991, a covert action campaign to "create the conditions for the removal of Saddam Hussein from power" was begun, and the CIA was given carte blanche to accomplish this (p. 59). In practice, this entailed organizing, arming, and training Iraqi groups that were favorable to the US and opposed Saddam. The rationale for this support to the rebels is clear. If a revolt's success means no longer having to pay the costs of containment or war, then it is entirely in the US interest to take measures to increase its likelihood.

In terms of the model, the conditions were ideal for the US willingness to incite a revolt. First, to

the extent that undermining Saddam's military and WMD programs is another benefit of sanctions, the net price of applying sanctions severe enough to incite revolt will be cheaper. Second, the costs of internal conflict in Iraq were declining over time, as Iraq's economy sank. And third, the US was doing its best to raise the chances of success for those groups it supported. According to Proposition 2, these three factors should all have contributed to the US willingness to impose sanctions and incite such revolts.

For its part, the first Clinton administration was initially wary of the risks of covert action to unseat Saddam, though it soon accepted them. But it never questioned the value of the sanctions on Iraq. These played a crucial role in undermining Saddam's military power and particularly his WMD programs, but they also provided the principal incentives for revolt. The partial relaxation of the sanctions agreed in May 1996 under the Oil-for-Food program simply traded off some of the costs of containment—it was necessary to counter growing international discontent with Iraq's impoverishment (p. 74).

In fact, no less than eight serious attempts to unseat Saddam were made during the years of containment, and the CIA was only known to be involved with two of them.[30] Moreover, at least four episodes of open civil conflict occurred, and these were understood to be caused by Iraq's worsening economic situation.[31]

The problem is that all eight attempts failed. In each and every case, Saddam managed to dodge the assassination attempt (though there were close calls), detect the coup early on and foil it (though several involved remarkably high-level conspiracy), and crush the revolt. Each time, the US had to revise downward its estimate of the likelihood of a successful revolt in the foreseeable future.

However, after 1996, when the international consensus on containment was beginning to show serious fractures, the steady-state costs of containment rose rapidly (p. 82–91). Thus, as Proposition 2 predicts, the US continued to support the opposition even as the evidence mounted that its chances of success were low (p. 96). As late as September 1998, two years after Saddam had

[30]There were coup, revolt, or assassination attempts in May 1991 (p. 59), June 1992 (p. 59), December 1993 (p. 68), January 1994 (p. 68), March 1995 (p. 72), May 1995 (p. 75), June 1996 (p. 80), and January 1999 (p. 93). The CIA was involved with the associated groups in the cases of March 1995 and June 1996.

[31]In 1993 (p. 68), 1994 (p. 68), 1995 (p. 75), and 1999 (p. 93).

broken or ejected all of the CIA-supported groups within Iraq, the US Congress passed the Iraq Liberation Act calling explicitly for Saddam's overthrow and providing the unprecedented sum of almost $100 million for that purpose. And in 1999, the Clinton administration began trying to resuscitate and re-organize the Iraqi opposition (p. 97).

In time, events such as the Kosovo War, a renewed attempt at Middle East peace negotiations, and the upcoming US election intervened. Given the exceedingly low chances of success, the outgoing Clinton administration shelved aggressive plans for pursuing internal regime change, and the Iraqi opposition fell apart (p. 98–100, 102).

After eight years spent trying to unseat Saddam, and little in the way of competent opposition left to support, the US gave up on inciting regime change in Iraq. Once it had done so, paying the costs of the still-high severity of sanctions no longer made sense, so the US agreed to the further relaxation of economic sanctions known informally as "Oil-for-Stuff," though it was determined to maintain military sanctions and to demand a new inspections regime in order to preserve the vise grip sanctions had provided on Saddam's WMD programs.

Thus, by the time the Bush (Jr.) administration entered office, there was no option other than indefinite containment or war. By 2002, the Director of Central Intelligence reportedly estimated the chance of success of a new program to overthrow Saddam at 10 to 20 percent (p. 290). Given the disarray of the opposition after 2000, and the historical record of the CIA's attempts with well-organized, well-trained groups, is is hard to believe that this was not a severe overestimate.

The model presented here thus explains why the US would maintain severe sanctions over the course of containment, despite their costs, and why it was (at least while a competent opposition existed) reasonable to expect that this would eventually lead to Saddam's overthrow. Indeed, many serious attempts were made, so in some sense the policy was successful. Though its estimate of the chances of success in this endeavor steadily ratcheted downward in the face of many failed attempts, as the costs of containment rose, it became willing to bet on long shots. Eventually, the plausible avenues for internal regime change were exhausted, and the costly sanctions most essential to encouraging revolt were dropped. By the end of containment, the only remaining option for being rid of Saddam was a US invasion.

5 Predation and the American War of Independence

The third and final source of costs in peace that I will analyze is *predation*: the expropriation of value from one actor by another. This expropriation may be costly because it undermines the victim's incentive to produce value. This idea is a very old one in economics, and is central to the economic analysis of taxation, growth, and many other phenomena. And there is a long line of research following Hirshleifer (1991) that explores the relationship among predation, inefficiency, and conflict. However, to my knowledge, the model presented below is the first to make explicit the connection between the costs of predation and the choice to fight a war, as opposed to peaceful (but potentially coercive) bargaining.

Predation occurs when one actor (the "producer") has a comparative advantage in production, while the other (the "predator") has a comparative advantage in coercion. That is, the producer is good at creating wealth while the predator is good at fighting wars. The predator can use his strength to expropriate, under threat of violence, some of the producer's wealth. The problem with this taxation is that it lessens the value the producer gets from his hard work, and so discourages production, or at least production that the predator can steal. This reduces the value that is available for either actor to consume, and so makes peace costly. If instead the two fought, then the producer would have a chance to end the predator's taxation, and the predator would have a chance to directly control the producer's effort. If the costs of peace exceed those of war, then they will fight.

I will explore this source of costly peace with a model that is tailored to the interaction between the colonies that later became the United States (henceforth, the "Colonists") and Britain, in the period following the Seven Years' War. This model is different from most bargaining models of war in that it incorporates an economy that is slightly more realistic than the usual "pie of size 1." Most importantly, what the two actors bargain or fight over is endogenously determined, rather than being fixed by assumption. While this complicates the analysis somewhat, it does seem a promising approach for developing more realistic bargaining models that can speak to the vast literature on the relationship between economic phenomena and war.

I then use the model to provide an account of the American War of Independence (henceforth,

the "war" or "Revolution"). The Seven Years War, which eliminated the common interest the Colonists and Britain had in cooperating to oppose France, exposed conflicting interests over the governance of the colonies that ultimately boiled down to how to divide the fruit of the Colonists' labor. Britain attempted to impose a series of unprecedented taxes, and when the Colonists resisted these taxes, Britain responded by initiating efforts to increase its control over colonial governments. The Colonists came to believe that peace would mean tolerating escalating taxation and ceding more direct control to Britain. Because the colonial economy was particularly responsive to taxation, peace would be very costly, and so the costs of peace came to be viewed as exceeding those of war and the Revolution occurred.

This account draws very heavily on Rackove, Rutten and Weingast (2000) and de Figueiredo Jr., Rackove and Weingast (2006), which to my knowledge constitute the only other rationalist account of the Revolution. These companion papers (henceforth "FFRW") report the results of a study that did much of the hard work involved in formulating a rationalist account of the Revolution, especially in developing a rationalist interpretation of the deep role that ideas about governance played in the conflict and in providing the evidence to support this interpretation. In fact, I have little to add to their understanding of the issues in contention between the belligerents.

What I can explain, using the model analyzed here and the evidence FRRW gathered, is why war happened rather than coercive bargaining over the future tax treatment of the colonies. This is a question that cannot be answered with the model in de Figueiredo Jr., Rackove and Weingast (2006), because that model does not allow bargaining—the Colonists must either accede fully to Britain's demands or fight. This formulation rules out the possibility that the Colonists and Britain could compromise on some division of policy-setting power between the two. By contrast, the model I present allows for such compromises but explains why they would be abandoned for war.

I'll begin with the setup of the model, then state the results and discuss the intuition, and then provide an account of the war.

5.1 A Model of Predation

Suppose we have two actors, Britain (indexed by B) and the Colonists (indexed by C). These actors are embedded in an economy in which there are two kinds of activity. "Transferable" activity produces things that are rivalrous in consumption and can be transferred between the two actors. That is, each product can be consumed by either, but not both, actors. Transferable products include most goods and services traded in licit markets. In contrast, "non-transferable" activity produces things which are rivalrous in consumption but which cannot be transferred between the two actors. In particular, only the actor that made some non-transferable product can consume it. For example, the standard non-transferable activity in economics is leisure: its production requires some allotment of a worker's resources (e.g., time), and once produced it cannot be moved to another actor. But goods produced and sold outside of the reach of the tax authorities, such as those trafficked locally on the black market, or in foreign trade via smuggling, or outside of authority's remit on the frontier, are also non-transferable.

The economy has a total endowment of resources $R \in \mathbb{R}^+$, which can be allocated to either transferable or non-transferable activity. Assume that the Colonists control the entire endowment R; in other words, C is the producer. This means that whatever Britain consumes must be transferred from the Colonists, and hence can only be drawn from the Colonists' transferable activity. In other words, B is the predator. Of course in reality Britain had productive resources of its own, but these were not under contention historically, so I do not include them in the model. Note that it would make no qualitative difference to the results if we instead assumed that the endowment was initially partitioned between the two players, but it would unnecessarily complicate the exposition of the model.

While Britain derives utility solely from transferable activity, the Colonists derive utility from both transferable and non-transferable products. For tractability we will assume that the Colonists' utility function is additively separable: it is the sum of the utilities derived from each kind of product. Since the two players will bargain over the disposition of the transferable products, we assume for simplicity that their utilities are risk-neutral over these. These assumptions allow us to

write their utilities as:

$$u_C(r,\tau) \equiv (1-\tau)(R-r) + l(r) \tag{4}$$

$$u_B(r,\tau) \equiv \tau(R-r) \tag{5}$$

Here, $\tau \in [0,1]$ is the proportion of transferable products that the predator consumes, r is the amount of resources devoted to non-transferable production, $R-r$ is the amount given to transferable production, and $l(\cdot)$ is the utility of consuming non-transferable products. We assume that $l(r) > 0$, $l'(r) > 0$, $l''(r) < 0$, and $l'''(r) > 0$, for $r \in (0,R)$, and $l(0) = 0$, $l'(0) = 1$, and $l'(R) = 0$. These are standard assumptions from economic theory that ensure that: C likes consuming non-transferable products; more consumption is better than less, but there are diminishing returns; C has decreasing absolute risk aversion; and C's optimal allocation between transferable and non-transferable production will always involve producing at least a little of both whenever the tax rate is neither 0 nor 1.

The sequence of moves is as follows. First, the Colonists make an offer to Britain or start a war.[32] If C makes an offer, B either accepts this offer or starts a war. If an offer is accepted, the producer chooses an allocation of resources between transferable and non-transferable production, production occurs, the agreed transfer is made, all products are consumed, and the game ends.

The Colonists' offer is to transfer a certain fraction ($\tau \in [0,1]$) of whatever transferable products are made to Britain. In economic terms, C has to choose a linear tax on transferable products, the proceeds from which will be given to B. This is an important assumption. It rules out non-linear tax schemes, in which the marginal rate of tax might depend on how much is produced. Allowing non-linear taxes that vary in the amount produced would not change the qualitative results, which require only that the tax scheme distorts C's allocation between transferable and non-transferable products and hence introduces inefficiency (i.e., costs) into peace, something that any such scheme would do. I discard such taxes for mathematical convenience.

However, a non-linear tax scheme that does not vary in the amount produced—a lump-sum

[32] It might be more realistic to assume that Britain was making offers to the Colonists, or that they were exchanging offers. However, this bargaining protocol simplifies the exposition, and the results do not depend on it, as whether war occurs does not depend on the allocation of the bargaining surplus.

tax—would not distort the producer's allocation and so would not make peace costly. The problem with such taxes, as economists well know, is that they create moral hazard. The tax collector (here, the predator) wants to assign taxes to an activity that represent some fraction of its full realizable value, but it often will not know what that value is, even in aggregate. If the participants in that activity anticipate that a lump-sum tax will be assessed, they have an incentive to realize less than its full value, so as to reduce the amount of tax that will be assigned. This creates inefficient distortion just as in the case of taxes that do vary in the amount produced. In principle, the tax collector could permanently assign the tax at the beginning, and promise not to alter it. If the participants believed this promise, there would be no more distortions and the full value would be realized. The problem here is that if this full value were above the tax collector's estimate, he would have an incentive to renege on his promise and raise the tax, which is precisely why participants wouldn't believe that promise and the distortions would continue. In other words, lump-sum taxes would end up varying with the amount produced in practice. I discard them and focus on linear taxes in favor of the simplicity of the latter, but the main result derived from a model in which lump-sum taxes and moral hazard were included would be qualitatively the same as the one given here. All that matters is how costly the resulting distortions are.

War is modeled as a costly lottery, won by the Colonists with probability p and Britain with $1-p$, with costs $d_C, d_B > 0$ regardless of outcome. If the Colonists win, then there will be no transfers of any kind to Britain. If Britain wins, then it is assumed to take direct control over the Colonists, their resources, and the tax rate, at some cost $s \geq 0$, which is the cost of exercising this control. I assume that $s \ll R$, so that control doesn't eliminate a large portion of the total value. In effect, the Colonists fight to be free of Britain's predation, and Britain fights to control the colonial economy.

In this formulation, the only thing the two players bargain over is the tax rate. We have presumed that it is not possible for Britain to take control over any portion of the Colonists' allocation of resources, except by victory in war. In theory, rather than imposing a tax, Britain could enslave some fraction of the colonial population, expropriate a portion of their capital, and occupy part of their land. While it would thereby eliminate the economic distortions that arise

from a tax on *products*, this predation on the *factors* of production generates other distortions, in decisions to invest in and expand these factors over time. A one-time transfer of a portion of the Colonists' factors to Britain won't satisfy the latter, because as the Colonists create new resources, Britain would also demand a share of these. But this is equivalent to what I model here, in that Britain imposes a tax (now on factors instead of products) which distorts incentives for creating value. Whether we are talking about the disposition of transferable products or that of productive factors, the point is that predation can distort incentives for production or investment and thereby make peace costly. War solves this problem, because if the Colonists win, there is no further predation and thus no distortion, and if Britain wins, it gains control of *all* of the Colonists' economic activity and is able to direct both the creation and employment of resources, at a cost.

Furthermore, we assume that the Colonists would not commit themselves to any allocation of resources, other than the one that is optimal under the agreed tax rate. Essentially, the only way the colonial governments can force their constituents to choose an allocation other than the one that is best given the tax rate is to exert direct control over them, just as Britain would if it won a war. However, doing so would leave these constituents powerless and unable to secure any utility for themselves in the face of predation by their own government. They would no more acquiesce peacefully to their own government doing this than they would to Britain's attempt.

Finally, all parameters of the game are assumed to be common knowledge.

5.2 Analysis of the Model

The analysis of the model aims to answer just one question. Under what conditions will the tax (predation) necessary to satisfy Britain make peace so costly that the Colonists would prefer to revolt? I will state the answer in a proposition and then discuss the intuition for it. The proof is in the appendix.

Proposition 3. *War will occur if and only if the sensitivity of the Colonists' allocation between transferable and non-transferable activity to taxation is high enough. When the costs of war are lower, the threshold sensitivity that separates war and peace is also lower.*

In this model, the only way that peace is efficient is if Britain agrees to a tax rate of zero. Any

positive tax rate encourages the Colonists to shift resources into non-transferable activity, which produces less value per unit of resource than transferable activity. As the tax rises, Colonists shift more resources into even less productive non-transferable activity, and the costs of these distortions grows at an accelerating rate. Thus, given that a tax rate of zero would not satisfy Britain, peace will be costly.

Given that there is a tax rate that would appease Britain, whether war or predation occurs depends only on which has the higher costs. If the destruction of war and the cost of Britain exerting direct control over the colonial economy if it wins exceed the distortionary costs of the lowest tax rate that would appease Britain, then the Colonists will tolerate predation. But if the reverse is true, then the Colonists will seek to throw off Britain's predation, violently if necessary, and Britain will fight to take direct control of the economy.

Whether war or predation is more costly depends principally on how sensitive the Colonists' allocation of resources between transferable and non-transferable activity is to taxation. If even large tax rates would only lead to small shifts in allocation, then the distortion caused by the tax is small, the Colonists continue to produce close to the full realizable value of their resources, and there is a tax that both Britain and the Colonists prefer to war. If, on the other hand, even modest tax rates lead to large shifts in the Colonists' allocation, then the distortion caused by taxation is large and the Colonists will produce much less than the full realizable value of their resources. Then, the tax that is necessary to satisfy Britain will have to be higher in order to account for the faster reduction in total value, and even the least such tax will not appease the Colonists.

Equilibrium here is inefficient whether it involves predation or war. The fundamental drivers of this inefficiency are an interlocking set of commitment problems and (unmodeled) asymmetric information. First consider lump-sum taxes. As discussed previously, in theory these would not create any distortions, because they don't alter the value of a given allocation on the part of the Colonists relative to any other. If Britain could commit to an initial assignment of these taxes, the Colonists would have no reason to alter their allocation from the efficient one, and the surplus from the lack of distortion makes possible a set of lump-sum taxes that both would strictly prefer either to varying taxes or war.

However, Britain cannot commit not to adjust these taxes in response to unexpectedly high value creation by the Colonists. So, if in setting the initial lump-sum taxes, Britain underestimates their productivity, then the Colonists will realize less than the full value so as to avoid revealing higher productivity and being subjected to a tax hike, and peace will be costly. This problem could be avoided if there were not asymmetric information about the Colonists' productivity and incentives to misrepresent it.

Even if Britain acknowledged this problem and switched to using taxes that varied in the amount produced (e.g., the linear taxes modeled here), this doesn't suffice to ensure that peace is costly. If the Colonists could commit themselves to the efficient allocation, then even varying taxes would not create costly distortions and both players could be made strictly better off than war or inefficient predation. They might be able to commit to this by having their government, or Britain, take direct control of their allocation decision. However, they would only cede this control if their government or Britain could commit not to take advantage of this control by taking all their transferable products. If either could commit, they would, but of course they cannot.

Thus, the presence of commitment problems and asymmetric information ensures that equilibrium must be inefficient, whether it involves distortionary predation or war. I turn now to applying these results to the American Revolution.

5.3 Stylized Facts of British-Colonial Relations

Here, I will first describe a set of stylized facts about the interaction between Britain and the Colonists in the 18th century. The evidence and argumentation for these facts is collected and presented in Rackove, Rutten and Weingast (2000) and de Figueiredo Jr., Rackove and Weingast (2006) ("FRRW"). I then argue that the model analyzed here explains why disputes over principles of taxation could lead to a large, expensive war. This section closes with a consideration of the alternative explanations for the war given by FFRW.

I will make use of four stylized facts that FFRW substantiate in their account, enumerated and described below. All that I have added to their arguments is a summary of the dispute between Ireland and Britain over the former's legal status, which undoubtedly provided a recent historical

analogy the Colonists could use to develop expectations about the ramifications of Britain's new policies toward them.[33] Readers seeking fuller documentation of the facts given below are referred to FFRW. For each, I also provide an interpretation of the fact in terms of the model presented here.

1. Before 1763, disputes over imperial policy (and authority) arose, but were always resolved peacefully, even though the underlying conflict was already present. There were three reasons for this. First, life for the Colonists was hard up to the early 18th century, so that there was relatively little surplus for Britain to extract. Second, the Colonists composed the vast majority of Britain's empire, and imperial policy-making was therefore attuned to their needs. Third, and most importantly, France was also powerful in North America, and formed an ever-present and very dangerous threat to Britain's colonies there. In opposing France, the Colonists and Britain had a very strong common interest, and the presence of the French in North America would have made open conflict between the Colonists and Britain very expensive because it would have rendered them vulnerable.

In the various wars fought with France and its Native American allies, the Colonists contributed substantially to their own defense, including quartering Britain's troops and raising their own militias to fight alongside them. They also paid great costs through the Navigation Acts, intended to halt colonial trade with France, though these also led to enormous smuggling.

Nonetheless, there were serious disputes over colonial policy during this era. In particular, Britain's Board of Trade sought reforms that would allow it to assert greater control over various colonies, in part to clamp down on rampant smuggling. These were protested by the Colonists, and failed to pass Parliament, presumably in order to maintain a unified front against France.

In the model, the restrictions of the Navigation Act can be thought of as implicit taxes on the Colonists, because they effectively transferred value from the restricted traders to Britain by undermining the traders' business but also hurting France. However, as long as France remained powerful in North America, these restrictions also improved the general well-being of the Colonists,

[33]This summary is drawn from the entry on "Declaratory Act, 1766" in the Gale Encyclopedia of US History, published by Cengage in 2007.

because they too benefited from the effort to contain France. In other words, these implicit taxes were effectively used to provide public goods for Britain and the Colonists. Thus, both *governments* had good reason to tolerate them, though individual colonists might not. The fact that many traders nonetheless shifted resources into non-taxable activity—smuggling—in such high volume indicates a high sensitivity on the part of the Colonists to taxation.

2. Trends in place before 1763 exacerbated the underlying conflict of interests, and laid the groundwork for its eventual exposure. First, the 1707 Acts of Union melded Scotland and England into Great Britain by uniting their Parliaments, but left Ireland out. Soon after, Britain passed the Declaratory Act of 1719, subordinating Ireland to both the Crown and Parliament of Britain, which could freely revoke Ireland's own Parliament, and asserting control over all legal actions to be taken in Ireland. The resulting abuse of Irish interests motivated Drapier's letters (written pseudonymously by Jonathan Swift), which paralleled the later pamphleteering of American radicals like Thomas Paine, and considerable unrest in Ireland.

Second, the Great Awakening of the 1730s and 40s greatly strengthened religious observance and religious diversity among the Colonists, and increased their distance from the Church of England, which at that time was reasserting itself in Britain. This raised the value of the religious liberties entailed in the colonists' self-government, liberties that were considerably greater than those enjoyed in Britain.

Third, the Colonists were rapidly getting richer. Their population expanded enormously over the 18th century, and the Colonists swiftly expanded frontier settlements, putting new lands into production and developing the continent's vast natural resources. The net result was that both overall wealth, and income per capita, grew substantially over this period.

Fourth, and most importantly, over the course of the 18th century, the long series of wars among the major powers, fought both in Europe and in the areas of colonization, resulted in radical increases in the power, extent, and indebtedness of Britain's empire. Securing this empire was immensely expensive and would require a much more capable and assertive imperial bureaucracy than had existed theretofore. In particular, to maintain its dominance, Britain would have to draw on the wealth of its colonies.

Of special import to the Colonists was Britain's victory in the Seven Years' War. The settlement of the war included the expulsion of the French from the bulk of North America, and largely ended any near-term threat to the Colonists from France. (It was precisely for this reason that it was safe for the rebelling Colonists to seek France's help during the war.) In doing so, it eliminated the strong interest the Colonists had previously shared with Britain in opposing France, and greatly reduced the costs of open conflict between the Colonists and Britain, for both.

In the model, the trends toward religious diversity and increased wealth among the Colonists can be thought of as increasing the value (R) that is potentially available to be "taxed," while Britain's new indebtedness increased the value to it of obtaining new sources of revenue. Britain might impose restrictions on religious practice, which would harm (transfer value from) colonial practitioners but benefit Britain in the form of the Church of England, or it might imposes taxes on the Colonists' new wealth. The treatment of Ireland after the Declaratory Act provided an example of what might happen to the Colonists if they allowed Britain to exert direct control over their affairs. But most importantly, the elimination of the French threat meant that the implicit (and high) taxes of the Navigation Acts were no longer providing public goods—they no longer benefitted the Colonists.

3. Britain well understood the risk of a colonial rebellion following the Seven Years' War, but it also needed money to pay for empire, had the power to demand it, and might need much more in the future. It therefore set out to reform colonial administration with an eye to its future needs. Thus began a series of attempts to impose unprecedented taxes on the Colonists. These measures were explicitly designed to accommodate the Colonists to new forms of taxation: the levies were very small; the taxed activity was chosen to be less visible in order to minimize popular aggravation and less evadable in order to minimize distortionary costs. Presumably, the intent was to prepare the Colonists for later increases in these taxes as Britain required them.

However, each new attempt brought unrest, evasion, resistance, and boycotts of British goods from the Colonists. Britain initially reacted by repealing one measure and passing a new one, hoping to find something the Colonists would accept. When this failed, it began to enforce compliance. Britain responded to resistance by punishing those leading the resistance and asserting more direct

control over colonial governance. Its punitive responses included suspending the assemblies of New York and Massachusetts, and closing the port of Boston and later blockading all of Massachusetts. Its assertions of control included taking over payment of colonial officials' salaries, and establishing a new customs board with British staff to enforce trade restrictions and new courts ruled by British appointees without juries to try smugglers.

In particular, the Declaratory Act of 1766, passed along with the repeal of the Stamp Act, held ominous portents for the Colonists' future. It was patterned nearly word-for-word after the earlier Declaratory Act for Ireland, discussed above. The Colonists understood exactly where this might lead. Thomas Jefferson, for instance, in response to the later Coercive Acts, argued that Britain's actions "too plainly prove[d] a deliberate and systematical plan of reducing us to slavery," which is the status the Colonists perceived Ireland to have.

Also, the Quebec Act of 1774 legally enshrined Catholicism in the newly acquired French-Canadian territories and expanded the remit of those territories into areas of the frontier adjacent to the Colonists'. It thus implicitly threatened the religious liberty of the Colonists, and also showed that Britain was willing to impose substantial harm on the Colonists' prospects for growth in order to satisfy its other interests of placating the French-Canadians and the Native Americans.

In sum, the days of conciliation in order to preserve unity against France were over.

In the model, Britain's gingerly attempts at increasing taxation can be thought of as trying to find ways to minimize the Colonists' subsequent transfer of resources into non-taxable activity. Britain's efforts to enforce compliance were attuned to the need to exert more direct control over the colonial economy, by clamping down on the local assemblies and making colonial officials directly responsible to Britain. These were initial steps toward the more thorough-going solution Britain presumably would have implemented had it won the war later on, and the new Declaratory Act made this intention clear. Finally, the real wealth of the Colonists lay not in their current prosperity, but in the vast opportunities for development represented by the territory beyond the frontier. By curtailing these opportunities, the Quebec Act effectively transferred a large amount of value from the Colonists to Britain (in the form of a different province of the Empire). It thus made clear that Britain would not limit itself to the small direct taxes (e.g., on sugar and tea) it tried to enact, but

would also impose much larger (though implicit) taxes.

4. The core of these disputes was not the actual taxes imposed, which were modest overall, but the principle that legitimated them. Both sides' material interests were inextricably entwined with the ideas of the Enlightenment and of the Imperial Era that Britain was entering, concerning proper forms of governance and definitions of sovereignty. The Colonists supported a form of imperial federalism they viewed as the earlier status quo, in which the Colonists had control over all internal affairs, political (including the selection of local legislatures), economic (including taxation and expansion), and social (including religious freedoms), while Britain had control over all foreign affairs, especially trade and security. In this view, freedom from tyranny lay in local governance and the principle of no taxation without representation.

By contrast, Britain supported an absolute form of parliamentary sovereignty: Parliament itself was the main bulwark against tyranny, and was therefore "unlimited and unlimitable." The ancillary doctrine of virtual representation, developed in part in response to colonial agitation, held that Parliament provided legitimate representation even to those (such as the Colonists) who had no role in the selection of its members.

Of course, each side's ideology was entirely consistent with its material interests. Sovereign Parliament could freely and rightfully extract wealth from the Colonists to pay the expenses of empire; self-governing Colonists could freely and rightfully enjoy their own wealth. The real taxes contested by the two sides were not the piddling measures Britain tried to enact, but the much larger taxes that might follow these under whichever principle reined supreme.

The model provides only the crudest representation of the ideological debate between Britain and the Colonists, in the form of bargaining over the implemented tax. However, this stylized fact makes clear that taxes—interpreted broadly to mean measures that transfer value from the Colonists to Britain and distort colonial affairs—were indeed the heart of the issue. In bargaining over taxes, the Colonists were concerned not with Britain's initial efforts at predation but with where the process might end: in effect, the final agreed tax.

These stylized facts reveal how the conflict of interests between Britain and the Colonists took shape. They explain why relations between the two were relatively pacific for most of the 18th century, and why these issues came explosively to the fore after 1763. They also render a clear rationalist interpretation of the disputes over principles of governance that connects these principles to implications for the future prosperity of the Colonists and predation upon them, and thereby resolves the mystery of why the Colonists would be provoked to the point of violence by the present, small taxes Britain tried to impose. However, despite all this, they do not make clear why the war happened.

5.4 Explaining the War Itself

Why couldn't the Colonists and Britain find a suitable compromise? After all, it's not as if Britain was unwilling to accomodate the principle of no taxation without representation. Earlier in the century it had unified with Scotland by combining the Scots Parliament with the English one. And it's not as though the Colonists were unwilling to compromise on a strong central government: not long after the Revolution, the weak Confederation was replaced with a much stronger United States government. Britain could have added some colonial representatives to Parliament, in accordance with the power and importance of the Colonists, in exchange for which the latter could have agreed to Parliament's sovereignty. Or, the two sides could have agreed to retain imperial federalism, with the Colonists governing their internal affairs, but with a specified set of activities that Britain would be free to tax. Yet neither of these, nor any other reasonable alternative, were given serious consideration by either side.

The model presented here explains why. Any settlement Britain would agree to would have to transfer enough value from the Colonists to at least equal the value Britain would expect to get from a war to reassert its sovereignty. And any settlement that appeased Britain would have to leave enough for the Colonists to meet the value they would expect to get from a war of independence. But transferring enough to satisfy Britain imposed costs on the colonial economy, in the form of resources shifted out of transferable activity. These costs were high, and the anticipated costs of war were low, so no mutually satisfactory compromises existed and the Revolution occurred.

Obviously, the validity of this account turns on whether the costs of peace, in the form of the distortionary costs of the taxation that Britain was expected to eventually demand, actually exceeded the costs of war. There are two good reasons to believe that they did.

First, the Colonists were adept at escaping the reach of authority. The colonial governments had nothing like the sophisticated imperial bureaucracy that was then coming into being in Britain. Settlers expanded the frontier, whether permission was given or not. The frontier also provided a ready escape for anyone fleeing the authorities. Traders routinely disregarded customs law and smuggled in volume, evading the implicit taxes of the Navigation Acts with gusto. The prevalence of these behaviors implies that the Colonists could quite easily shift their productive efforts into untaxed activity when needed. This in turn implies that they would respond sensitively to more taxation, and thus that the costs of peace in the near term would be high.

Second, and more importantly, the vast majority of colonial prosperity lay in the future, something the Colonists well knew. The expansion of settlement, the clearing of new land for agricultural and later industrial production, the extraction of natural resources, and the development of canals and other infrastructure to support the movement of goods, were the keys to the colonies' wealth. And yet at the time of the Revolution, all of these had only just begun. This fact has important implications for the costs of both war and peace.

With regard to the costs of war, it implied there was much less to be destroyed in a war begun at the time. War might be expensive in the near term, and indeed it was. But war posed no threat to the colonies' future wealth, as it was not even fought in the territory that contained most of this wealth, and the settled areas had much less capital and infrastructure in existence than they would later. Thus, at worst a war could destroy a small fraction of the colonies' total value, meaning that the costs of war for both Britain and the Colonists would be modest in the long view.

With regard to the costs of peace, it implied that the long-term sensitivity of the colonial economy to taxation was very high. Realizing the wealth tied up in America's interior would require immense entrepreneurial effort that heavy taxation might discourage. Since any negative impact on this growth would compound exponentially over time, even a slight decline in the natural growth rate would be extremely costly. Indeed, this presumably explains the very low tax burden

the Colonists imposed once they had won their independence and maintained until at least the Civil War.

Certainly, the Colonists could anticipate that Britain's policies would affect the growth rate. The Quebec Act demonstrated Britain's willingness to hand the Colonists' future growth opportunities over to more acquiescent provinces. Moreover, Britain continually hounded the colonial authorities to limit territorial expansion, largely because of its need to avoid a substantial, ongoing commitment of military resources to fighting wars with Native Americans on the frontier. This suggested that, in the future, colonial expansion would be delayed as necessary to bring peace to the frontier and enable Britain to respond to urgent military needs elsewhere. Given the experience of Britain's many wars in the first half of the 18th century, these delays could occur regularly.

This is not to say that, had Britain retained the colonies of the future United States, it would not have invested heavily in their development. This development was in Britain's interest, too, and indeed it was partly why Britain had a policy of "salutary neglect" toward the colonies in earlier times. However, if the colonies remained in the Empire, this interest would sometimes be traded off against Britain's other interests. In particular, the Colonists could expect to be regularly squeezed to help pay for Britain's many wars. The problem was that, after the defeat of France in the Seven Years' War, these wars were not relevant to the Colonists' well-being and thus constituted a pure transfer that could only distort the colonial economy.

The bottom line is that, both now and in the future, the Colonists had good reason to anticipate that the taxation that would be needed to satisfy Britain would impose great costs on their own prosperity. Greater even than the costs of a war, now, to win independence and end future predation, which could only become more expensive over time. So, they declared independence. For its part, Britain could anticipate that the Colonists would resist its extraction of revenue, and had already begun to increase its control over the colonies. Faced with a declaration of independence, Britain chose to fight a war to take full control and end resistance to its taxes.

5.5 Alternative Explanations

FFRW give two subtly different explanations for the war in their two papers. I will describe each in turn, and then discuss the fundamental flaw that is common to both. The problem is that the reasoning underlying these explanations rules out bargaining between the Colonists and Britain. This leads to war, but for empirically implausible reasons, as I will explain.

First, the explanation in Rackove, Rutten and Weingast (2000) relies on a commitment problem. In this story, the stability of Britain's governance of the Colonists is built on the continuation of a long-standing policy of benign neglect. This policy is interpreted by the Colonists as an endorsement of their philosophy of imperial federalism, itself derived from the ideas developed in the time of Britain's internal struggle over self-government, and so is satisfactory to them. However, changes in Britain's interests, deriving mainly from its war victories and the concomitant expansion of its empire, lead Britain to take steps that depart from tradition and violate imperial federalism, though they are fully consistent with Britain's actual philosophy of parliamentary sovereignty. These steps break tradition and thereby undermine the Colonists' confidence in Britain's commitment to imperial federalism. Thus, the Colonists worry that Britain cannot commit not to intervene willy-nilly in their internal affairs, something they could not tolerate. They then initiate a war to secure their independence.

The second explanation, in de Figueiredo Jr., Rackove and Weingast (2006), offers a similar story, but is focused more on understanding how it could (rationally) be true that both the Colonists and Britain were each surprised by the vehemence of the other's reaction to what they viewed as rightful, uncontroversial behavior. In this telling, the Colonists were unaware that Britain's philosophy had shifted from imperial federalism to absolute parliamentary sovereignty over the course of the 18th century, and remained so as long as the common interest of opposing France led Britain to treat the Colonists lightly. The Colonists thus believed that they played a kind of repeated prisoner's dilemma with Britain, in which mutual cooperation consisted of the Colonists' loyalty to the empire and Britain's non-intervention in their internal affairs. When the French threat was removed, Britain's interests changed and it imposed modest, though novel, taxes that were in keeping with its philosophy and that it thus still saw as cooperative. Because the Colonists

incorrectly believe that Britain still supports imperial federalism, they view this as (surprising) defection and so respond with defection: a declaration of independence. And because Britain is unaware that the Colonists have false beliefs, it is surprised by their defection and also responds with defection: war.

Both of these explanations for the war are seriously flawed, and for the same reason: they do not allow the two sides to bargain over a new arrangement for governance. In both explanations, the choice of this arrangement is binary. Britain either intervenes or it doesn't. In the first explanation, the commitment problem is not the one to which Rackove, Rutten and Weingast (2000) refers. If the only two choices are intervention or not, and the intervention is anticipated to be substantial, then Britain would strictly prefer a war to no intervention, and the Colonists would strictly prefer a war (the result of declaring independence, given Britain's preference) to intervention. Neither would commit to an option other than the one it chose, because the other option would leave it strictly worse off. War happens because the disputed issue is completely indivisible. This is still war due to a commitment problem: the two sides would prefer to flip a coin to determine the outcome, but they can't commit to respect the result.

In the second explanation, war happens because of asymmetric information about the game being played and about Britain's preferences. However, there is no incentive to misrepresent this information on the part of either side. Thus, if they had the opportunity to communicate their understandings in the course of bargaining over a settlement, they would no longer have false beliefs about the other's interests, and the only thing standing in the way of peace would be the binary choice of settlement.

It seems highly implausible to assume that only two settlements were available to the Colonists and Britain. What would stop them from considering a compromise, in which Britain would be given some authority over the Colonists (e.g., a specified set of activity that could be taxed, or specified revenue or rates of tax), and the Colonists would retain some autonomy (e.g., over religious affairs and frontier policy)? If an agreement existed that both sides preferred to war, then they could overcome initially incorrect beliefs about the other's interests in the course of bargaining. They would then be able to identify this agreement (or another like it), and its implementation

would be stable (within the bargaining range) because each side could credibly threaten war if the other violated it. Thus, FRRW cannot explain why the war itself occurred.[34]

The model I have presented in this section can: no such agreement existed, not because it was not physically implementable or because of incorrect beliefs, but because the anticipated distortionary costs of the taxation necessary to satisfy Britain reduced the value of the peace to the point that war, which would eliminate these costs, was preferred.

6 Developing the New Explanation

The principal intent of this essay was to convince the reader that costly peace is an empirically relevant rationalist explanation for war. To that end, it analyzed three sources of costly peace—arming, imposition, and predation—and showed how, in each case, underlying commitment problems or asymmetric information forced actors into inefficient equilibria, and war occurred whenever its costs were lower than those of peacefully coexisting with arming, imposition, and predation. It used these three sources to provide analytic narratives of three different cases: the Iraq War, the civil conflict in Iraq between the wars with the United States, and the American War of Independence. Thus, costly peace is empirically relevant: it explains at least one inter-state war, one intra-state war, and one extra-state war better than the available alternatives.

It is my hope that scholars of war will take from this the main implication: costly peace should be added to the rationalist toolkit for explaining the origins of wars. It is useful for understanding the wars studied here, and it might be useful for understanding many more. A second implication is that, even if there is some war for which scholars have confidently identified the presence of a commitment problem, asymmetric information, or both, these do not suffice to explain the occurrence of the war. To have a complete account of its causes, it must also be established that costly measures *other* than war, which the belligerents could have taken to address the identified drivers of inefficiency, either did not exist, were insufficiently effective to suppress the underlying driver, or were simply more costly than war.

[34] More generally, two-by-two models of conflict such as the prisoner's dilemma are, by construction, unable to illuminate why two actors would choose war over a peaceful settlement.

For those interested in using the bargaining theory of war to make informed predictions about the war-proneness of some dyad, or to perform statistical analysis of the historical record of wars, there is a third implication. Analysts must beware of using only indicators of either commitment problems or asymmetric information as predictors, without also taking account of the possible presence of measures available to the players that, while costly, might be cheaper than war and hence prevent war from occurring. Ignoring these other measures will bias the resulting predictions toward war.

Finally, for those who seek to advise policy-makers on how to respond to the threat, initiation, or continuation of war around the world, there is a fourth implication. Wars that occur due to costly peace are not like those that occur in the absence of costly peace—some interventions which would improve the lot of the (potential) belligerents in the absence of costly peace will actually harm them in its presence. Preventing or stopping such a war may simply force the belligerents to suffer even greater costs than those of the war. A decision to prevent war or impose peace should only be made if there is some belief that the war is not being caused by costly peace.

For both scholars and practitioners, then, it is valuable to be able to distinguish which wars occur due to costly peace. Unfortunately, and whatever the merits or flaws of the analytic narratives I presented in this essay, divining the principal cause of any war is most certainly a black art. That said, its value is such that I believe scholars ought to do more of it, and in the most rigorous and thorough fashion possible. To that end, I will conclude this essay with a brief set of suggestions for how to go about identifying other wars that are plausibly due to costly peace, and a speculative list of some possible candidates for such wars.

6.1 How to Distinguish War due to Costly Peace Empirically

In practice, there are at least three criteria that should be met in order to distinguish whether a war can plausibly be explained with reference to the costs of peace. These are not exhaustive, but they are in some sense minimal requirements, and each corresponds to a potential pitfall for a scholar eager to attribute some war to costly peace.

First, the participants had to have the means to overcome any commitment problems or asym-

metric information without going to war. This is potentially a very challenging requirement, because it is hard to think of a war that does not involve at least *some* movement in the balance of power and *some* asymmetry of information. And, since most wars are preceded by substantial spending on arms, it is also difficult to think of a war in which costly peace played *no* role in its causes. When more than one of these explanations appears to contribute substantially to the occurrence of a particular war, then it is very hard to identify one or another as being the dominant cause. To have any confidence in an attribution of a war to costly peace, there must be a reasonable case that, had the participants not gone to war, peace would not entail much shifting power or asymmetric information.

Second, peace must actually be substantially costly *overall*. It is not enough that one actor sees peace as costly because he does not get his way, or because he suspected the other actor was taking advantage of him, or because of any reason that he alone expected to gain from war. Both actors must anticipate that peace between them will yield substantially less than the total value it could, were it not for some source of costs.

Third, and this is one is the most tricky, war must somehow, in expectation, *reduce* the costs of peace. This criterion is much harder to meet than it might first appear. A story must be told for why there was no deal the actors could agree to, no mechanism they could implement, that would mitigate the source of costs in peace, other than war. In other words, how exactly does war solve the problem of these costs, and why can't this solution be implemented peacefully?

6.2 Other Wars Due to Costly Peace

I make no claim that the three sources of costly peace identified here—arming, imposition, and predation—are exhaustive empirically. They are only the ones I have thought of so far, and there are probably more. But, in addition to the particular wars discussed in previous sections, there are a number of other wars that might be attributed to these three sources. It seems likely that analyzing these wars through the lens of costly peace would be a valuable exercise.

Arming: Wars of consolidation, secession, and succession might derive substantially from arming costs. Fearon (1995) offered the possession of a given throne as a possible example of an indivisible

issue, but then noted that, while in principle a throne is divisible—it can be shared or alternately held, or its concomitant territory divided, by two potential occupants—in practice, the norms of monarchy make it appear indivisible. An alternative explanation that either supplants norms or provides an implicit explanation for where the norms come from is that splitting up territory or sharing a throne requires that both occupants maintain the militaries necessary to preserve their claims. If there are economies of scale in the maintenance of such militaries or in the protection of one's territory from outside powers, then the added costs of multiple, smaller armies may exceed the costs of war to retain a unified throne. Thus the costs of arming prevents territory from being divided too far below a certain efficient scale, and when this scale is reached, war will be preferred to further division.

Imposition: Could some of the civil wars that occurred during the Cold War have been due to the imposition of penalties and rewards by the two superpowers? In the global competition for states not solidly in either superpower's camp, each superpower routinely supplied substantial military and other aid to its favored elements within such states, which was intended to increase the leverage of this group relative to the other favored by the opposing superpower. If the peaceful coexistence of the two groups implies continuing costs for the superpowers of arming and support, then both superpowers may try to incite their clients to fight a war to consolidate control over the state and eliminate these costs. The same can be said more generally for the desire of hegemonic states to avoid bearing the costs of war to eliminate some nuisance by encouraging a local party to do it for them.

Predation: Wars over raiding and piracy—essentially, land- and sea-borne predation—have occurred since antiquity, and continue in the rangelands of Africa and potentially off the coast of Somalia. Do the costs of predation cause these wars? If the producers fight to end the predation, do the predators fight to temporarily erode the producer's defenses and so take full control over its resources while its strength recovers? Also, many wars have been fought over prohibition, from the international opium wars of the 19th century to more recent intra-state wars in Colombia and Mexico. Because the social costs of (certain) drugs' use are thought to exceed the individual benefits,

drug traffickers effectively prey on productive society. Is this what causes these wars?

Proofs

Proof of Proposition 1. We proceed by backward induction to find subgame perfect equilibria (SPE). Bayesian solution concepts are unnecessary because all equilibria will be in pure strategies, and any US uncertainty about Iraq's efforts to obtain nuclear weapons will actually be irrelevant to the outcome.

First suppose that, at the beginning of a period, Iraq obtains or already has nuclear weapons. It is a standard result in games of this type that the unique SPE is for the US to make the minimum satisfactory offer to Iraq, equal to its per-period payoff from war, or $1 - p^n - d^n_{IR}$ (Bas and Coe, 2011).[35] Thus the values of this nuclear subgame are $V^n_{US} = \frac{p^n + d^n_{IR}}{1-\delta}$ and $V^n_{IR} = \frac{1-p^n-d^n_{IR}}{1-\delta}$ for the respective players.

In the previous subgame, in which Iraq does not have nuclear weapons, there are essentially five possibilities for the equilibrium. First, the US could refrain from containment, while conceding as little as possible to Iraq, but enough that Iraq would prefer this concession to war; call this "toleration." Second, the US and Iraq might agree to a cooperative equilibrium in which the US did not contain but offered a concession to Iraq in exchange for its agreement not to pursue nuclear weapons, and threatened to punish detected pursuit with war; call this "cooperation without containment." Third, the US could engage in containment, while giving Iraq as little as possible but enough to avoid war; call this "containment." Fourth, with the US engaging in containment, the two countries might agree to a cooperative equilibrium just like in the second possibility; call this "cooperation with containment." Finally, the US could go to war.

The method of proof is to calculate the continuation values associated with each of these five possibilities and show that the calibrating assumptions rule out toleration and cooperation with or without containment. This leaves just containment or war as possible equilibria, and we will see that which exists depends only on a comparison of the costs of each.

Toleration: Clearly, the US should offer Iraq as little as possible while avoiding war; offering

[35] I assume here that Iraq can and will credibly signal to the US that it has nuclear weapons as soon as it obtains them (e.g., with a test detonation). It has clear incentives to do so, because nuclear weapons improve its war value and so necessitate a more generous offer from the US if peace is to occur, and the US would only be willing to make this offer if Iraq revealed it had the weapons—otherwise, Iraq could be bluffing.

any more would be unnecessary and leave the US worse off. Under this strategy, Iraq would strictly prefer to try to get nuclear weapons because it expects to do strictly better in the subgame in which it has the weapons. This implies that it may not be possible for the US to make small enough offers to hold Iraq to its war value. If the US simply offered nothing at all, Iraq would still get:

$$V_{IR}^t = 0 + \delta \left[\lambda V_{IR}^n + (1-\lambda) V_{IR}^t\right] \tag{6}$$

$$\Rightarrow V_{IR}^t = \frac{\delta\lambda}{1-\delta(1-\lambda)} V_{IR}^n \tag{7}$$

Simplifying, this value will be at least equal to Iraq's war value whenever:

$$1 - p - d_{IR} \leq \frac{\delta\lambda}{1-\delta}(p - p^n + d_{IR} - d_{IR}^n) \tag{8}$$

The first calibrating assumption is equivalent to this inequality.

The US value of toleration is the total value of the game, minus Iraq's value from toleration, or $V_{US}^t \equiv \frac{1}{1-\delta} - V_{IR}^t$. By contrast, the value from war, if the US had first engaged in containment, is $W_{US}^c \equiv \frac{p^c - d_{US}^c}{1-\delta} - c_{US}$. The second calibrating assumption is equivalent to $W_{US}^c > V_{US}^t$, or:

$$1 - p^c + d_{US}^c + c_{US}[1 - \delta(1-\lambda)] + \frac{\delta\lambda}{1-\delta}(d_{US}^c + d_{IR}^n) < \frac{\delta\lambda}{1-\delta}(p^c - p^n) \tag{9}$$

Since the US strictly prefers war to toleration, the latter cannot be in equilibrium.

Cooperation without containment: Without containment, the best the US could do to dissuade Iraq from pursuing the weapons would be to offer the most generous possible settlement, conditional on the discontinuation of Iraq's nuclear program, and then threaten to respond to any detected resumption of this program by initiating containment and going to war. The value of such a deal for Iraq must be no more than $\frac{1}{1-\delta} - W_{US}^c$, because if it were any more than this the US would strictly prefer war over the deal. Suppose the candidate deal achieves this maximum; we will show that Iraq would still have an incentive to cheat under a condition equivalent to the third calibrating assumption.

Let $W_{IR}^c \equiv \frac{1-p^c-d_{IR}^c}{1-\delta} - c_{IR}$ be Iraq's value from war when the US is engaging in containment.

If Iraq tried to get nuclear weapons under this best deal, its continuation value (starting from the next period) would be:

$$\lambda V_{IR}^{n} + (1-\lambda)(1-\sigma)\left[\frac{1}{1-\delta} - W_{US}^{c}\right] + (1-\lambda)\sigma W_{IR}^{c} \tag{10}$$

Here, the first term is the value Iraq gets if it acquires nuclear weapons, the second is the value it gets if its effort is unsuccessful but goes undetected, and the third is its value if its effort is unsuccessful and detected, leading the US to start a war. If Iraq's value of cheating exceeds the value of the deal, then at every period, Iraq has an incentive to renege on the deal. Setting the value of cheating to be greater than the value of the deal and simplifying, we obtain:

$$(1-\lambda)\sigma\left[\frac{d_{US}^{c} + d_{IR}^{c}}{1-\delta} + c_{US} + c_{IR}\right] < \lambda\left[\frac{p^{c} - p^{n} - d_{US}^{c} - d_{IR}^{n}}{1-\delta} - c_{US}\right] \tag{11}$$

The third calibrating assumption is equivalent to this condition. Thus, the US would not offer any deal without containment, because it would expect Iraq to renege on it.

Containment: Just as in toleration, the US should offer Iraq as little as possible while avoiding war, and Iraq would then strictly prefer to try to get nuclear weapons. Similarly, it may not be possible for the US to make small enough offers to hold Iraq to its war value, so Iraq's continuation value under containment must be the higher of its war value and the value of getting nothing until it obtained nuclear weapons, or $V_{IR}^{c} \equiv \max\left\{W_{IR}^{c}, \frac{-c_{IR}+\delta\lambda^{c}V_{IR}^{n}}{1-\delta(1-\lambda^{c})}\right\}$.

The fourth calibrating assumption is equivalent to $W_{IR}^{c} \geq \frac{-c_{IR}+\delta\lambda^{c}V_{IR}^{n}}{1-\delta(1-\lambda^{c})}$: in other words, the first term in the maximum at least equals the second. Under this assumption the US value from containment is just $V_{US}^{c} \equiv \frac{1-c_{US}-c_{IR}}{1-\delta} - V_{IR}^{c}$.

Cooperation with containment: There is a deal under containment that the US could offer that would cause Iraq to forego nuclear weapons, but the US won't offer it in equilibrium. The fourth calibrating assumption implies that the US is squeezing everything it possibly can out of Iraq, even accounting for the fact that Iraq is pursuing nuclear weapons. The only way to get Iraq to stop is to make it an offer generous enough that it is left willing to eschew nuclear weapons, given that if it tries to get them and is detected the US will punish it by war or containment, both of

which give Iraq the same value of W_{IR}^c. The minimum such offer will have a value for Iraq exactly equal to the value of a deviation from this deal to pursuing nuclear weapons:

$$V_{IR}^{cc} \equiv \lambda^c V_{IR}^n + (1-\lambda^c)\sigma^c W_{IR}^c + (1-\lambda^c)(1-\sigma^c)V_{IR}^{cc} \tag{12}$$

$$\Rightarrow V_{IR}^{cc} = \frac{\lambda^c V_{IR}^n + (1-\lambda^c)\sigma^c W_{IR}^c}{\lambda^c + (1-\lambda^c)\sigma^c} \tag{13}$$

Notice that if $\lambda^c > 0$, then $V_{IR}^{cc} > W_{IR}^c$, because $W_{IR}^c < V_{IR}^n = W_{IR}^n$ by assumption. Thus, the least generous deal that would cause Iraq to prefer (weakly) to refrain from nuclear weapons development would leave Iraq with more than its war value. But then the value of this deal for the US is $V_{US}^{cc} \equiv \frac{1-c_{US}-c_{IR}}{1-\delta} - V_{IR}^{cc} < \frac{1-c_{US}-c_{IR}}{1-\delta} - W_{IR}^c = V_{US}^c$, so that the US strictly prefers not to make this more generous offer. So cooperation with containment is not in equilibrium.

War: It was assumed in the setup of the model that the US would prefer to engage in containment before going to war because of the military advantages containment provides, so the value of war for the US is just W_{US}^c.

Since toleration and cooperation with or without containment are ruled out by the calibrating assumptions, equilibrium must be whichever of war or containment gives a higher value for the US. Setting $W_{US}^c > V_{US}^c$ and simplifying results in $d_{US}^c + d_{IR}^c < \delta(c_{US} + c_{IR})$. This establishes the proposition. □

Proof of Lemma 1. Assume that the US has already moved and imposed sanctions of severity α and demanded a settlement (inclusively) to the left of q_α. We will determine the conditions under which a deal exists that Saddam and the rebels would prefer to civil war.

First notice that, if $\alpha = 0$, so that no sanction is imposed, the internal actors can safely ignore the US demand and will implement the settlement $q = p + d_R$, as this gives the rebels just enough to prefer peace to war and leaves Saddam with all the surplus from avoiding the costs of war, so that he would strictly prefer this offer over any other and over war (which would only give Saddam $p - d_S$). Now, the only reason the US would pay the cost of imposing sanctions is to get a more desirable outcome than would occur in the absence of sanctions, so in equilibrium, if $\alpha \neq 0$, then the US demand q_α must be less than $p + d_R$. That is, by imposing sanctions, the US aims to coerce the

internal actors (particularly Saddam) to choose a settlement closer to 0 than would occur absent US interference.

This implies that Saddam effectively has at most three options. First, he can start an internal war with the opposition. If he loses, he receives $0 - d_S$, as the rebels implement their most favored policies, but if he wins, his payoff depends on whether he would then set policies so as to meet the US demand. In equilibrium he would do so if it gave at least as high a value as spurning the US demand. Thus, his war payoff is $w_S \equiv p \cdot \max\{q_\alpha, 1-\alpha\} - d_S$. Second, Saddam could offer the opposition the minimum necessary to avoid war. If this minimum does not meet the US demand, then it is the value of the internal game under sanctions, $1-\alpha$, minus the rebels' war payoff, $w_R \equiv 1 - p - d_R$, or $p + d_R - \alpha$. If this minimum does meet the US demand, then it is just q_α, since we know this is less than $p + d_R$ and so will satisfy the rebels. Third, Saddam could offer the minimum necessary to meet the US demand (if this is different from the second option), meaning an offer and payoff of q_α. Any other offer Saddam might make would either yield the same payoff as one of these choices, or would leave Saddam strictly worse off than at least one of these three, because it would entail giving the opposition more than was necessary to appease either them or the US. In equilibrium, Saddam would choose whichever of these offers gave him the highest payoff.

This implies that the US can always incite an internal war, by ensuring that the costs of peace exceed the costs of war for Saddam's regime and the rebels. To do so, the US must first choose a demand that is to the left of Saddam's payoff from war, or $q_\alpha < w_S$. This implies that $w_S = p(1-\alpha) - d_S$, and ensures that Saddam would strictly prefer internal war to meeting the US demand. And second, the US must impose sufficiently severe sanctions that Saddam would strictly prefer internal war to giving the opposition enough to avoid it:

$$w_S = p(1-\alpha) - d_S > p + d_R - \alpha \tag{14}$$

$$\Leftrightarrow (1-p)\alpha > d_S + d_R \tag{15}$$

The left side of the inequality is the cost of peace: the rebels forego the opportunity to fight and, if they win, to eliminate the sanctions. The right side is the cost of internal war. Conversely, if the

US imposes less severe sanctions than this, then Saddam would prefer appeasing the opposition over war, and if the US demand is less stringent than Saddam's war payoff, then he will always prefer meeting it to war. □

Proof of Proposition 2. The method of proof is to calculate the values for the US of its optimal war-inciting sanctions regime and of its optimal peaceful sanctions regime, and then determine a sufficient condition for the former to exceed the latter.

Given that the US chose to incite a war within Iraq, it should impose the cheapest sanction severity that would do so, because imposing more severe sanctions would be more costly and bring no benefit. Using Lemma 1, the value to the US of internal war is $w_{US} \equiv (1-p) + p(0 - c_{US}) - d_{US} - s\left(\frac{d_S + d_R}{1-p}\right)$.

To find the optimal peaceful sanctions regime, observe from the proof of Lemma 1 that if α and q_α result in peace, then the implemented outcome will be the higher of $p + d_R - \alpha$ and q_α. Because of this, it cannot be that $q_\alpha > p + d_R - \alpha$ in equilibrium. If this inequality held, the US could make a more stringent demand (i.e., costlessly reduce q_α) and Saddam would still meet it, making the outcome more favorable to the US. So in a peaceful equilibrium, Saddam will offer $p + d_R - \alpha$, and the US value of this sanctions regime will be $v_{US}(\alpha) = 1 - p - d_R + \alpha - s(\alpha) - c_{US}$. The first-order condition with respect to α is $s'(\alpha^*) = 1$, and the assumptions on $s(\cdot)$ imply that a unique solution for α^* exists and that it is the global maximum.

Thus, the US value from the optimal peaceful sanctions regime is less than that of a war-inciting regime whenever:

$$v_{US}(\alpha^*) < w_{US} \tag{16}$$

$$\Leftrightarrow -d_R + \alpha^* - s(\alpha^*) < (1-p)c_{US} - d_{US} - s\left(\frac{d_S + d_R}{1-p}\right) \tag{17}$$

The left side does not depend on c_{US}, but the right side grows unboundedly in it, so that incitement is better for the US if c_{US} is high enough. For the converse, first observe that, by virtue of the assumptions on $s(\cdot)$, it must be that $\alpha^* \gg s(\alpha^*) > 0$. If d_R is small (assume this), then the left side of the inequality is always non-negative. The right side, however, is negative for small enough

c_{US}. The other statements of the proposition follow immediately from the inequality. □

Proof of Proposition 3. The method is backward induction to find SPE. First I determine the Colonists' optimal allocation for a given tax rate and compute the players' war values and peace values under the tax. The occurrence of peace is equivalent to the existence of a tax rate that gives both players at least their war values, so I find the conditions for this.

The Colonists choose $r \in [0, R]$ to maximize $u_C(r, \tau)$. The first-order condition is $1 - \tau = l'(r^*)$. The assumptions on $l(\cdot)$ ensure that this equation has a unique solution for r^* that is the global maximum, so the optimal allocation can be written as a function of τ, or $r^*(\tau)$. Note that the assumption that $l'(0) = 1$ implies that $r^*(0) = 0$, so that the Colonists would devote all their resources to transferable production in the absence of taxation.

Let the sensitivity of the Colonists' allocation decision to taxation at a given tax rate τ be defined as the derivative of r^* with respect to τ, which by implicit differentiation of the first-order condition is $-1/l''(r^*(\tau))$. When this is small (large), the Colonists' allocation shifts slowly (rapidly) in τ.

If Britain won a war, and thereby had the ability to choose both the allocation of resources (at a cost s) and the tax rate, it is obviously best for it to allocate all resources to transferable production, and to transfer all of this production to itself. If the Colonists won the war, the tax rate would be zero and the allocation choice would be $r^*(0) = 0$. Thus the players' war values are:

$$W_C = p \cdot [(1-0)(R - r^*(0)) + l(r^*(0))] + (1-p) \cdot [(1-1)(R-0) + l(0)] - c_C = p\,[R] - c_C \quad (18)$$

$$W_B = p \cdot [0(R - r^*(0))] + (1-p) \cdot [1(R-0) - s] - c_B = (1-p)\,[R - s] - c_B \quad (19)$$

Now, the value the players obtain from peace, and its derivative with respect to τ is:

$$P(\tau) \equiv u_C(r^*(\tau), \tau) + u_B(r^*(\tau), \tau) = R - r^*(\tau) + l(r^*(\tau)) \quad (20)$$

$$P'(\tau) = -r^{*'}(\tau) + l'(r^*(\tau)) \cdot r^{*'}(\tau) = \frac{\tau}{l''(r^*(\tau))} \quad (21)$$

So $P'(\tau) < 0$ for $\tau > 0$ and peace is efficient if and only if $\tau = 0$.

This implies that peace will be least costly if Britain is given a tax rate just high enough to meet his war value, which in turn maximizes the value left for appeasing the Colonists. If it exists, this rate is found by equating Britain's war value to its peace value and finding the lowest solution for the implied peacetime tax rate $\underline{\tau}$:

$$(1-p)\left[R-s\right]-c_B=\underline{\tau}(R-r^*(\underline{\tau})) \tag{22}$$

Because Britain's peace value is continuous in τ and equal to zero for $\tau=0$, if there is any τ that would satisfy it, then there is a $\underline{\tau}$ that would *just* satisfy it. Thus, if there is no solution for $\underline{\tau}$, then there is no tax rate that would satisfy the predator and there must be war. If the sensitivity of the Colonists' allocation to taxation is approximately zero, then the right side of the equation approaches R as $\underline{\tau}$ approaches 1, and the Intermediate Value Theorem guarantees the equation has a solution. If instead the sensitivity is very high, then r^* rapidly approaches R as $\underline{\tau}$ increases from zero, so that the right side does not rise much above zero and thus the equation has no solution. Also, the existence of a solution is less likely if p, c_B, and/or s are low.

Suppose now that there is a solution to the equation, and let the lowest such solution be $\underline{\tau}^*$. The only thing that remains is to check whether the Colonists prefer war to offering $\underline{\tau}^*$. Because this tax equates the Britain's peace and war values, the Colonists will prefer war if and only if:

$$P(\underline{\tau}^*)\leq W_C+W_B \tag{23}$$

$$\Leftrightarrow R-r^*(\underline{\tau}^*)+l(r^*(\underline{\tau}^*))\leq p[R]+(1-p)[R-s]-c_C-c_B \tag{24}$$

$$\Leftrightarrow c_{war}\equiv(1-p)s+c_C+c_B\leq P(0)-P(\underline{\tau}^*)\equiv c(\underline{\tau}^*) \tag{25}$$

The left side of this inequality is the expected costs of war: the destruction the two sides will suffer regardless of who wins, plus the cost of Britain exerting control over the colonial economy weighted by its chance of victory. The right side is the costs of peace at the least tax rate that would satisfy Britain. Clearly, the inequality is easier to satisfy if p is high and/or s, c_C, and c_B are low.

The left side does not depend on the sensitivity of the Colonists' allocation to tax and is positive. If the sensitivity approaches zero, then $c(\underline{\tau}^*)$ also approaches zero and the inequality is violated,

meaning the Colonists prefer peace. As the sensitivity grows, $\underline{\tau}^*$ rises and $c(\underline{\tau}^*)$ increases. For high enough sensitivity, either there will no longer be a solution for $\underline{\tau}^*$ or the inequality will be satisfied. This establishes the proposition. □

References

Baliga, Sandeep and Tomas Sjöström. 2008. "Strategic Ambiguity and Arms Proliferation." *Journal of Political Economy* 116(6):1023–1057.

Bas, Muhammet A. and Andrew J. Coe. 2011. "A Model of Proliferation and Prevention." Working Paper.

Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal and Barry R. Weingast. 1998. *Analytic Narratives*. Princeton University Press.

Bilmes, Linda and Joseph Stiglitz. 2006. "The Economic Costs of the Iraq War: An Appraisal Three Years after the Beginning of the Conflict." National Bureau of Economic Research Working Paper 12054.

Bush, George and Brent Scowcroft. 1999. *A World Transformed*. Vintage.

Cordesman, Anthony H. 2002. *Iraq's Military Capabilities in 2002: A Dynamic Net Assessment*. Center for Strategic and International Studies.

Davis, Steven J., Kevin M. Murphy and Robert H. Topel. 2006. "War in Iraq versus Containment." National Bureau of Economic Research Working Paper 12092.

de Figueiredo Jr., Rui J. P., Jack Rackove and Barry R. Weingast. 2006. "Rationality, Inaccurate Mental Models, and Self-Confirming Equilibrium." *Journal of Theoretical Politics* 18(4):384–415.

Debs, Alexandre and Nuno Monteiro. 2010. "Nothing to Fear but Fear Itself? Nuclear Proliferation and Preventive War." Working Paper.

Director of Central Intelligence. 2002. "Iraq's Continuing Programs for Weapons of Mass Destruction." National Intelligence Estimate, October.

Duelfer, Charles A. 2002*a*. Testimony before the Foreign Relations Committee of the United States Senate, July 31.

Duelfer, Charles A. 2002*b*. "Weapons of Mass Destruction Programs in Iraq." Testimony before the Subcommittee on Emerging Threats and Capabilities, Armed Services Committee of the United States Senate, February 27.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Fearon, James D. 2011. "A Simple Political Economy of Relations among Democracies and Autocracies." Working Paper.

Flibbert, Andrew. 2006. "The Road to Baghdad: Ideas and Intellectuals in Explanations of the Iraq War." *Security Studies* 15(2):310–352.

Hirshleifer, Jack. 1991. "The Paradox of Power." *Economics & Politics* 3(3):177–200.

Iraq Intelligence Commission. 2005. "Report of the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction." Unclassified Version, March 31.

Iraq Survey Group. 2004. "Comprehensive Report of the Special Advisor to the DCI on Iraq's WMD." September 30.

Jackson, Matthew O. and Massimo Morelli. 2009. "The Reasons for Wars: An Updated Survey." Forthcoming in The Handbook on the Political Economy of War, edited by Chris Coyne, Elgar Publishing.

Krueger, Alan B. 2006. "The Cost of Invading Iraq: Imponderables Meet Uncertainties." *The New York Times* . March 30.

Krueger, Alan B. 2007. "Economics, Iraq, and the War on Terrorism." Summary of remarks at the National Economists Club, September 25, 2007.

Lake, David A. 2010. "Two Cheers for Bargaining Theory: Assessing Rationalist Explanations of the Iraq War." *International Security* 35(3):7–52.

Mazarr, Michael J. 1995. *North Korea and the Bomb: A Case Study in Nonproliferation*. Palgrave Macmillan.

McBride, Michael, Gary Milante and Stergios Skaperdas. 2011. "Peace and War with Endogenous State Capacity." *Journal of Conflict Resolution* 55(3):446.

McBride, Michael and Stergios Skaperdas. 2007. Explaining Conflict in Low-Income Countries: Incomplete Contracting in the Shadow of the Future. In *Institutions and Norms in Economic Development*, ed. Mark Gradstein and Kai A. Konrad. MIT Press pp. 141–162.

McBride, Michael and Stergios Skaperdas. 2009. "Conflict, Settlement, and the Shadow of the Future." Working Paper.

Mearsheimer, John J. and Stephen M. Walt. 2003. "An Unnecessary War." *Foreign Policy* (134):50–59.

Nordhaus, William D. 2002. The Economic Consequences of a War with Iraq. In *War with Iraq: Costs, Consequences, and Alternatives*. American Academy of Arts and Sciences.

Pollack, Kenneth M. 2002. *The Threatening Storm: The Case for Invading Iraq*. Random House Inc.

Powell, Robert. 1993. "Guns, Butter, and Anarchy." *American Political Science Review* 87(1):115–132.

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton University Press.

Powell, Robert. 2002. "Bargaining Theory and International Conflict." *Annual Review of Political Science* 5(1):1–30.

Powell, Robert. 2006. "War As a Commitment Problem." *International Organization* 60(1):169–203.

Powell, Robert. 2011. "Bargaining, Fighting, and State Consolidations." Working Paper.

Quester, George H. 2000. *Nuclear Monopoly.* Transaction Publishers.

Rackove, Jack N., Andrew R. Rutten and Barry R. Weingast. 2000. "Ideas, Interests, and Credible Commitments in the American Revolution." Working Paper.

Reiter, Dan. 2003. "Exploring the Bargaining Model of War." *Perspective on Politics* 1(1):27–43.

Silverstone, Scott A. 2007. *Preventive War and American Democracy.* Routledge.

Slantchev, Branislav L. 2010. "Borrowed Power: Debt Finance and the Resort to Arms." Working Paper.

Wallsten, Scott and Katrina Kosec. 2005. "The Economic Costs of the War in Iraq." AEI-Brookings Joint Center for Regulatory Studies Working Paper 05-19.