

Rationality, Inaccurate Mental Models, and Self-Confirming Equilibrium: A New Understanding of the American Revolution

Rui, de Figueiredo, Jack Rakove, and Barry R. Weingast*

August 2000

1. Introduction

Rational choice theorists have long been ambivalent about the rationality postulate. Although many agree that humans have imperfect foresight and inaccurate understandings of the world within which they act, no satisfactory approach relaxing the rationality postulates exists. Theorists have taken several related tacks in their attempts to relax the boundaries of rationality. One group of scholars have attempted to incorporate specific lessons from cognitive science, psychology, or experiments that reveal limits on rational thinking. Recent examples include Camerer (19**), Denzau and North (1994), Elster (1999); Frey (1997), Frank (19**, 19**); Kanemen and Tversky (1979), Lupia and McCubbins (1998), and Rabin (19**).¹ A second approach reflects scholars who argue that under

*Assistant Professor, Haas School of Business, Berkeley; Coe Professor, Department of History, and Professor, Department of Political Science, Stanford University; and Senior Fellow, Hoover Institution, and Ward C. Krebs Family Professor and Chair, Department of Political Science, Stanford University. The authors gratefully acknowledge helpful conversations with Colin Camerer, John Ferejohn, Robert Gibbons, Stephen Krasner, David Laitin, Mathew McCubbins, Paul Milgrom, James Morrow, Douglass North, Robert Powell, Andy Rutten, Thomas Sargent, and Robert Wilson.

¹This list hardly exhausts the categories of work on this topic. For example, there is the long-standing work of Simon (19**) and March and Simon (19**) on bounded rationality; game theorists have studied “finite state automata (see, e.g., **); and there exists several collections of essays exist where various proponents and critics of rational choice theory discuss these issues (see, e.g., Cook and Levi 1990, and Mansbridge 1990).

certain circumstances, particular mechanisms of interaction imply that cognitive limits are not binding (Clark, 1997a,b) and Satz and Ferejohn (1994); see also North (1999, ch**), drawing on Hayek (1960, ch 2). Third, a group of scholars using game theory and related approaches have explored alternative definitions of rationality that directly incorporate individuals limited knowledge and abilities (see Fudenberg and Kreps 1995, Fudenberg and Levine 1993,1998, Greenberg 1994, Kalai and Lehrer 1993, and Sargent 1999).

In this paper, we draw on this third literature. Rather than incorporating a specific type of limit on rationality, we develop a new method that allows us to model directly rational action when agents have incomplete or inaccurate mental models of their world. At times, individuals act in such great ignorance of their world that unexpected results occur at every turn. These circumstances are not the subject of this paper. We instead study circumstances where mental models have typically evolved over time and where they have, in a particular sense, stood the test of time. The perspective requires an explanation of how, in the face of their ignorance, rational individuals nonetheless continue to rely on their inaccurate mental models. Our answer is not of the form often given those drawing on cognitive science, such as the hypothesis that individuals have a tendency to disregard evidence that contradicts their own point of view. Instead, we incorporate follow the game theorists noted above by incorporating inaccuracies directly into the model.

In the real world, humans face a variety of cognitive limits. But this does not imply that they are irrational or that they ignore experience. Following Greenberg (1994) and Sargent (1999), we assume that people attempt to understand their world, and their theories about the world in equilibrium are consistent with their experience. Because the players are not omniscient and have limited experience, they may have incomplete or inadequate models of their environment.

In terms of the approach, we allow the players in a game to have two types of inadequate understandings of the world. First, they may not completely understand the game they are playing; and second, they may have mistaken beliefs about the strategies used by the other players. Although the players can have inaccurate models of their environment, these inaccuracies cannot be arbitrary. The approach constrains inaccuracies to occur within an equilibrium framework that imposes discipline on the types of mistakes and misunderstandings. The framework requires that nothing the players observe along the equilibrium path disconfirm their beliefs about the game they are playing or about the strategies used by others. Although mutual misunderstandings are possible, they must be mutually consistent in this equilibrium sense.

This approach contrasts with the typical non-rational accounts that merely posit that the actors misunderstand their world in particular ways. In these approaches, the errors can be arbitrary. In our approach, the errors must be consistent with experience, in particular, with observations of the other players' choices.

To study these phenomena, we employ the concept of self-confirming equilibrium, a generalization of the traditional Nash equilibrium framework in game theory (Fudenberg and Levine 1993, 1998, Greenberg 1994, Sargent 1999). In the context of extensive form games, self-confirming equilibrium relaxes the common knowledge assumption that all players have the same understanding of their strategic environment. Players instead may have very different understandings of their strategic situations. The equilibrium component imposes discipline by requiring that nothing an agent observes disconfirms her views about her environment. The operational difference between self-confirming equilibrium and Nash equilibrium is thus twofold: self-confirming equilibrium

allows the players to have certain types of incorrect conjectures about the structure of the game and about one another's strategies.

The self-confirming equilibrium framework allows us to model an environment where a player's predictions about one another's behavior are inaccurate. Indeed, a player may be ignorant of her uncertainty, a phenomena ruled out in the traditional Nash equilibrium framework. The framework also allows us to model an common aspect of disputes; *viz.*, that both sides believe they are acting reasonably while the other side is acting unreasonably. In the Nash equilibrium framework, such interpretations are either epiphenomenal or irrelevant.

The approach allows a type of surprise in which a player's mis-conjectures about another's behavior result in behavior that was not anticipated. Consider a set of players in a self-confirming equilibrium, perhaps one that has been stable for multiple generations. Suppose that an exogenous shock occurs which, given their strategies, causes one or more players to alter their preferred moves. Because these moves were off-the-path prior to the exogenous shock, the players had no prior experience with this environment. Hence their beliefs about one-another's strategy may have been inaccurate — though never invalidated. Under these circumstances, one or more players may surprise the others by making an unanticipated choice.

We apply our framework to the problem of the American Revolution, helping to resolve several unexplained puzzles in the historical literature. Historians of the Revolution emphasize the role of ideas in underpinning the revolutionary crisis.² These historians focus on the differences between how the Americans and the British interpreted the principal issues. The British and the

²Major works include: Morgan (1957), Bailyn (1967), Greene (1986), Rakove (1979, 1996), Reid (1995), and Wood (1969).

Americans profoundly disagreed about the nature of the constitution governing the exercise of political authority and policymaking. Specific issues included the nature of British authority within the colonies, the nature of sovereignty, and the structure of the empire. Also central to the dispute was a debate about the mechanisms protecting liberty in America.

From the rational choice stand point, a critical omission in the literature is that historians do not connect the realm of ideas with the realm of action. Indeed, historians fail to explain a range of questions about the American Revolution (see Rakove, Rutten, and Weingast 1999). For our purposes, we consider three puzzles. First, Historians do not adequately explain why either side fought about abstract ideals, and thus why the two sides failed to come to some accommodation despite their differences. Second, why did the colonists become so incensed over the seemingly trivial taxes imposed by the British after the Seven Years' War ended in 1763. Finally, if the clash of fundamental ideals led the two sides to armed conflict, how could the British and the Americans be unaware of their profound differences during the previous 100 plus years of cooperation? An explanation for the Revolution must account for what allowed the two sides to cooperate for 100 years despite their differences, only then explode in a few short years.

Along with our companion paper (Rakove, Rutten and Weingast 1999), we here provide new answers to these questions. We begin with the theory of credible commitments to interpret how the American theorists argued that the traditional institutions of the British empire credibly protected American rights. During the revolutionary crisis, Americans developed a theory about how a range of institutions provided credible commitment to their rights and liberties. Relevant institutions included the structure of the empire, the division of authority between colony and metropole, and the role of colonial assemblies in providing for a range of American liberties.

Using the concept of credible commitments, we provide the critical link between the ideals expressed by the Americans during the revolutionary debates and the realm of politics, economics, and everyday life. Alarm rose across the colonies as many Americans came to believe that the British were systematically dismantling these credible commitments, putting their ways of life at risk. Further, British interventions in the colonies deeply challenged the view that the British were benevolent (Rakove, Rutten, and Weingast 1999).

Using the self-confirming equilibrium perspective, our account helps interpret these claims and to address the above puzzles. Turning to the ideas themselves, the Americans looked back to the English ideas of the late seventeenth century to understand what they considered the constitution of the empire (Greene 1986, Reid 1995). According to this perspective, long-established practice and custom became enshrined with constitutional status. Americans also interpreted the empire as having a federal structure (though they did not use this label): the British had control over the system-wide policies, such as trade and security, while the Americans retained control over their domestic affairs, including taxation, religion, and commerce. This control also covered property rights in a range of assets, such as land and slaves. The interpretation of empire as federal allowed the Americans to conclude that sovereignty within the empire was divided — indeed, that it had been for over 100 years.

In combination, these two principles had a critical implication: The American's constitutional perspective implied that the empire's federal structure, enshrined in several generations of practice, had attained constitutional status. According to this interpretation, the British lacked authority to regulate or otherwise control domestic colonial affairs. Any attempt by the British to do so violated the constitution and was therefore tyranny. The tax imposed by the British may have been trivial in

itself, but if the British had authority to impose that tax, revolutionary theorists argued, they had authority to alter all colonial laws and rights.

The British interpreted these issues very differently. By the latter half of the eighteenth century, the British had begun to develop a new view of the constitution that matured in nineteenth century. This view held that Parliament was the supreme authority in Britain and that it had the power to interpret and alter the constitution. The British also believed that the structure of the empire was a matter of policy choice having no constitutional status. Allowing the colonists control over their domestic affairs was a privilege, not a right. The fact that for 100 years the British had chosen not to exercise their powers to regulate domestic colonial affairs in no way impaired their right to do so. Moreover, the British had no reason to believe that the colonists saw these issues any differently. The British believed that they had the discretion to take back or otherwise alter these privileges. Finally, in the British view, sovereignty could not be divided. Thus, if any one had it, they did, not the colonies.

As historians emphasize, the British and the Americans held mutually inconsistent views about their world — with an explosive potential. As the crisis unfolded, the Americans' constitutional interpretation led them to view the British intervention into their domestic affairs as tyranny. Whereas the British constitutional interpretation led them to view American resistance to their policies as treason. Each side considered itself as acting lawfully and legitimately while the other was wholly in the wrong. Consider the following quotes from two of the Masters. According to Morgan (1977,44), “Everything that followed seemed to confirm suspicions on both sides. Repeal of the Stamp Act momentarily lulled the colonists' fears, but the Townshend Acts renewed them, and the colonists' antagonism to the acts fanned English suspicion of a colonial drive toward

independence. Every move by either side seemed to bring the prophecies of the other side closer to fulfillment.” Similarly, Bailyn (1967,158-159) suggests:

“the beliefs and fears on one side of the Revolutionary controversy were as sincere as those expressed on the other. The result, as anticipated by Burke as early as 1769, was an ‘escalation’ of distrust towards a disastrous deadlock: ‘The Americans,’ Burke said, ‘have made a discovery, or think they have made one, that we mean to oppress them; we have made a discovery, or think we have made one, that they intend to rise in rebellion against us...we know not how to advance; they know not how to retreat...Some party must give way.’”

Our self-confirming equilibrium model helps explain and elucidate the two side’s contradictory interpretations, including how they led to revolution and war. First, we show that in the 100 years prior to the Seven Years’ War (1756-63), the British and the Americans were in a self-confirming equilibrium. Although the two sides had very different ideas about the structure of the empire and about British powers, prior to the war, the British had no incentive to exercise their powers. Thus, the occasion that would generate evidence of the conflicting interpretations never arose. Nothing about the two views came into conflict, so neither side obtained evidence from the other to indicate that their conjectures clashed with the other’s views or possible actions.

The Seven Years’ War represented a dramatic shock to the environment. Prior to the war, the pervasive French threat in part underpinned trans-Atlantic cooperation. Both sides needed each other and were willing to bear costs to maintain the mutually beneficial relationship. The removal of the French threat at the end of the Seven Years’ War altered the basic structure of the empire, including both sides’ willingness to cooperate. The British gained a clear interest in regulating domestic colonial affairs, in part as a means of controlling a much larger empire that now included French Canada and the large territories containing Indians hostile to the American colonists.

The new configuration of interests following the Seven Years’ War brought to the fore the contradictory views over ideals and the mechanisms protecting American rights. What was

previously off the path behavior for the British — intervening in the American domestic affairs — became on the path, much to the surprise of the Americans. To the British, this exercise of authority did not constitute defection from cooperation, and they had no reason to expect that the Americans would. The Americans, however, did interpret intervention as defection; in reaction they initiated a trigger strategy. To British eyes, American resistance appeared as a bid for independence, taking advantage of the absence of the French threat. American resistance thus appeared as a unilateral defection from cooperation, and the British also reacted by exercising a trigger strategy. Put simply, each side viewed the other as the first to defect from cooperation.

Paralleling the historians' arguments, the model reveals how each side had an inadequate mental model of the other. Yet, for over 100 years neither side obtained any evidence to the contrary — no evidence emerged to cause either side to believe that their views were inadequate and inaccurate. By altering both sides' incentives, the Seven Years' War first generated a surprise to each, resulting in an escalating conflict, revolution, and war.

We develop our argument as follows. Section 2 provides the background on the revolution, including an account of the historical literature, the role of ideas, and a short narrative of events. Section 3 describes the self-confirming equilibrium framework. Section 4 applies the framework to the American Revolution. The heart of the paper, this section shows how the American and British views could differ yet not cause conflict during the century prior to the Seven Years' War, only to explode into conflict after the war. Our conclusions follow.

2. Constitutional Ideals During the American Revolution, Including a Brief Narrative

To a large extent, the agenda of modern historians of the American Revolution remains influenced by the Progressive historians, such as Charles Beard (1913). The Progressives' simplistic notions of economics, economic self-interest, and politics led them to deprecate the role of ideas in the American Revolution. Reducing ideas to caricatures, the Progressives argued that the Revolution was a power grab by colonial elites, who used the notions of liberty, tyranny, constitution, and independence to solidify their rule at home. In the Progressives' view, ideas were at best epiphenomena and, at worst, tools used by men of means to manipulate others by co-opting and pacifying them.

Modern historians have shown beyond doubt that the Progressives' view cannot be sustained. Contrary to the Progressive account, the colonial ideology was consistent and coherent. It was also well grounded in existing British constitutional thought (Bailyn 1967, Greene, 1994, Morgan 1995, Rakove 1996, Reid 1995, and Wood 1969). These historians emphasize the centrality of ideas and ideology in the development of the American Revolution. Americans devoted considerable effort to theorizing about liberty, the constitution, sovereignty, and the theory of the empire. Indeed, the dominant view among historians is that the revolution took place because of differences in ideals. In the words of one prominent historian,

The Revolution becomes comprehensible only when the mental framework, the Whig world-view into which the Americans fitted the events of the 1760's and 1770's, is known. "It is the development of this view to the point of overwhelming persuasiveness to the majority of American leaders and the meaning this view gave to the events of the time, and not simply an accumulation of grievances," writes Bailyn, "that explains the origins of the American Revolution." (Wood 1966,169)

. . . If the origin of the American Revolution lay not in the usual passions and interests of men, wherein did it lay? . . . Never before in history had a people achieved a "revolution by reasoning" alone. [162] The Revolution was thus essentially intellectual and declaratory. . . [163]

The debates between the British and colonists over the ideals of constitutionalism, sovereignty, liberty were not simply debates about abstract ideals, however. Before drawing the link between these ideals and everyday politics and economics, we study the ideas' content.

Contrasting views of the constitution

To a large extent, both sides shared the same ideals. Both accepted that abuse of power by government was a crucial problem, and that it could best be controlled by rule of law. In what follows, we will discuss the ideas that came to dominate in the American Colonies and in Britain. This ignores considerable differences of opinion in each side of the Atlantic. In our companion paper, we explain how these ideas came to dominate in the colonies (Rakove, Rutten, and Weingast 1999).

Underlying the broad agreements on ideals, however, were deep disagreements over the exact nature of the institutionalization of rule of law, particularly as it applied to the empire. For example, the British and Americans disagreed about the substantive American rights protected by the British constitution and the institutions established to protect those rights. Moreover time, the two sides also came to disagree about each other's commitment to those ideals. Put crudely, the colonists worried that the new policies proposed by the British after the Seven Years' War signaled a change in the constitution of the empire. They feared that constitutions that could be easily changed provided little protection of their rights. In light of these changes, colonists sought assurances that the British would respect their liberty. At the same time, the British worried that colonial resistance to the new policies

signaled a change in colonial attitudes towards the empire. They worried that claims about fundamental rights could be changed easily, and thus provided a weak foundation for membership in the empire. In light of these changes, imperial officials wanted assurances that any concessions would not lead to American independence.

To the British, there was only one Constitution in the Empire, the British Constitution (Greene 1986,62). For them, governing the Empire was the same as governing Britain. By 1776, this meant that the British vision of the Imperial Constitution included many features usually associated with the mature, nineteenth century British Constitution. Most important of these was the notion of parliamentary sovereignty, the idea that no authority stood over Parliament. In practice, this meant that whatever Parliament did was, by definition, constitutional.³ In addition, the British had begun to articulate the notion of “virtual representation.” This idea — that Parliament represented all parts of the empire regardless of whether they had actual representatives in Parliament — rationalized a relatively narrow franchise even within Britain.

The colonists rejected the equation of the British and Imperial constitutions. Like the British, they believed that Parliament had authority over matters relating to empire-wide issues, notably, trade and security. Unlike the British, they believed that the colonies — through their local assemblies, courts, and an executive appointed by the crown — had authority over local matters, such as property rights, religion, social control, and taxation.⁴ They supported this view by appealing to the Whig tradition of the late seventeenth and early eighteenth centuries, a tradition that stressed

³Hayek (1960, 176), in his classic on liberty, referred to this notion as “the principle of parliamentary sovereignty unlimited and unlimitable.”

⁴ According to Wood (1966, 166-67, quoting Morgan 19**, p **), “As early as 1765, the Whigs “laid down the line on which Americans stood until they cut their connections with England. Consistently from 1765 to 1776 they denied the authority of Parliament to tax them externally or internally; consistently they affirmed their willingness to submit to whatever legislation Parliament would enact for the supervision of the empire as a whole.”

the role of custom and precedent as sources of constitutional legitimacy (Reid 1995). The colonists treated the long-standing division of power within the empire — with British control over the empire-wide issues and colonial control over all local matters — as constitutionally mandated. Put simply, colonists believed that the long-standing practice of sovereignty over local matters enshrined these powers with a constitutional status.

The American view. The colonists derived their view of the imperial constitution from their experience. From the start, the British colonies were unlike those of those of Spain or Portugal. Instead of building on military defeat of indigenous empires, the British relied on large scale settlement by adventurers, lured by the promise of a better life. Since that better life was not a sure thing — life was unusually harsh for the early settlers — the King had to offer various concessions to lure potential colonists. Among these were grants of local political power, including that of a local legislature.⁵

In the century prior to the revolutionary crisis, the *de facto* power of the colonial assemblies continued to grow. Although the King appointed most colonial executives, he was dependent in important ways on the colonial legislatures. This dependence reflected economic necessity: given the costliness of trans-Atlantic communication, such decentralization was the only viable alternative. Metropolitan control would simply have been too expensive to be effective.

At the same time that practice was evolving towards what Greene calls “negotiated authority,” that is meaningful decentralization, law was continuing to celebrate the form of centralized power. Thus, during the seventeenth and early eighteenth century, various statutes

⁵The absence of valuable treasure — a major factor determining the structure of the Spanish empire — made these concessions less costly for the English crown than for the Spanish crown (see North, Summerhill, and Weingast 1999).

provided the King with authority over the colonies. For instance, he continued to appoint most colonial governors, and the acts of all colonial assemblies were subject to review by the Privy Council (Smith, 1965). In practice, however, that authority was largely a dead letter. For example, governors were routinely charged to rule with the aid of the assemblies. The decline of imperial participation culminated in the Walpole administration's adoption of what Burke later called a policy of "wise and salutary neglect" during the 1730s. As a result, by 1750, the institutions charged with overseeing the details of colonial policy, such as the Board of Trade, had "atrophied" (Greene, 1991, p. 97). In short, with respect to their internal affairs, the colonies were self-governing.⁶

Of the political practices that had grown up to govern the empire, the colonial assemblies were at the heart of many Americans' concerns over liberty. Although Americans disagreed about large numbers of ideals and policy matters, most agreed that their legislature was central to protecting these ideals. Members of self-sufficient religious communities, such as those in rural Massachusetts, might be indifferent to British imperial trade regulation, but they believed that colonial legislatures were central to protecting their religious freedoms that distinguished America from Britain. Similarly, many traders might care far more about British imperial trade policy than any other issue, but they depended on their legislature to maintain low and stable taxes and to define and protect their property rights and wealth. Whatever Americans valued as their liberty, the colonial legislature was central to protecting it.⁷

⁶ In fact, as Weingast (1995) argues, the division of authority in the empire matched a type of federalism called market-preserving federalism.

⁷ Not all Americans agreed that legislatures were central to the protection of freedom. Many minorities, anticipating the argument of Federalist # 10, feared that legislatures would become tools by which majority factions would oppress minorities. For them, the British Empire provided a "controlling common umpire," without which the colonies would become "a field of blood, a scene of terror and desolation." (Potter, 1983, pp. 143, 35)

For the colonists, the fact that they had exercised authority through their local legislatures established a constitutional claim. They argued that the British Constitution was like the common law, a set of rules built on practice and precedent. Drawing on earlier Whig traditions, the colonists argued that custom and precedent were a much more effective way to restrain tyranny and arbitrary rule than procedure. In arguing for custom, they were not, as many modern scholars argue, indulging in a nostalgic attachment to the past. Rather, they saw custom as a commitment device, since it, unlike statutes, could not be changed by any small group, such as the Crown or Houses of Parliament.

From this perspective, the historic principles governing the empire, at least as interpreted by the Americans, gave them a degree of protection — indeed, of credible commitment — against the imperial government. The notion of divided sovereignty reflected, in effect, a view that the empire had a federal structure: Britain controlled truly empire-wide issues, such as trade and security; while Americans controlled local matters, such as economic and social organization, through their legislatures. As long as the Americans and the British held to this structure, the principal American values and freedoms — including their livelihoods, their property and their religious practice — were protected from arbitrary action by the British crown or Parliament.

From the colonial perspective, a constitution founded on precedent embodied what is now called a “bright line credible commitment” mechanism. The federal structure of the empire was in part arbitrary; there were many alternative ways to organize the empire. Yet this structure had the advantage that deviations were easy to observe (hence the “bright line”) and thus easy for both sides to police. The division of authority inherent in the federal system governing the British Empire served as a bright-line, a rule that was so clear that virtually any citizen could tell when it had been

breached. As coordinated reactions were the only hope for the American colonists to police the British, the bright line mechanism facilitated a coordinated reaction to breaches.

The British view. Although the British based their views of the imperial constitution on the same practices as the colonists, they drew very different lessons from those practices. The British came to see their particular institutions of governance, especially the King in Parliament, as the chief bulwark against tyranny. Over the eighteenth century, this approach led them closer to articulating the position of parliamentary sovereignty, under which any act of Parliament was, by definition, constitutional (Blackstone, 1765, 156).

From the British perspective, precedent and practice had very different constitutional standing than they did for the Americans. To the British, the governance structures of the Empire while of great value, did not have constitutional status. These structures were a matter of policy dictated by expediency and efficiency. As such, Parliament could alter the empire's governance structure whenever it wanted. This view implied that the British believed that the powers exercised by colonial assemblies were privileges granted by Parliament to the Americans, not constitutional rights restricting parliamentary authority. Finally, the British believed that sovereignty was indivisible, implying that, through its Parliament, Great Britain was sovereign over all aspects of the empire.

The constitution and the revolutionary crisis

Until the end of the Seven Years' War, American and British differences led to few lasting problems. As Greene (1994) documents, many disputes over authority arose during this period, as

colonists clashed with imperial officials on a wide range of issues. Before 1763, the metropole and colonies had managed to resolve these disputes peacefully.

Following the long series of wars from the late 1730s that ended with the Seven Years' War, British imperial policy underwent significant change, reflecting the changing demands of empire. Although the governance of the empire, American in particular, changed after 1763, the British believed the change fully within their constitutional authority. They had fought an expensive war in part to protect the colonies; the demands of a huge new empire required a change in governance. From the British perspective, the American demands and arguments seemed flagrant violations of the imperial system. Despite the Americans' claims of affirming the status quo, their demands appeared to the British as a call for revolutionary separation from the Empire (Tucker and Hendrickson 1982).

A short narrative of events

The British and the colonists reacted to the shock of the end of the Seven Years' War against this background (Christie 1966, Tucker and Hendrickson 1982). The end of the War and the removal of the French from North America had important implications for both the British and the Colonists. First, it decisively ended the French threat for the time being. Second, victory led to creation of a much larger, more secure, more diverse — and more deeply indebted — Empire. Third, the first two implied major changes in the interests of both Britain and America. Self-interest no longer underpinned a harmony of interests, in which a harm to one also harmed the other.

The Seven Years' War had several implications for American behavior. First, in the absence of the French threat, Americans would be less likely to suffer burdens, because there would no

concomitant gains. The same held for the British: the now much larger Empire implied that the original American colonies no longer loomed so large in the British imperial calculus. Standard theories of bargaining among self-interested parties suggests that the Americans would be less conciliatory to the British than before the war. Second, a contemporaneous economic change had occurred in America: By mid-century, for the first time, America had real wealth to expropriate (McCusker and Menard, 1985). This, too, had an implication for both the British and the Americans. For the British, it implied that for the first time there was something to extract from the Americans; for the Americans, it implied that they had more to fear from British opportunism.

At the beginning of the crisis, the radicals emerged to argue that the British were malevolent and out to destroy American liberty. Yet these arguments fell on deaf ears. These seemingly wild claims simply failed to match experience.

Over time, however, the British behavior came to take on a different cast. A series of acts seemed clearly to signal a change in British policy with respect to American sovereignty, the constitution, the empire, and the role of the British in regulating domestic colonial affairs. This began with the Sugar Act of 1764, explicitly designed to raise revenue, and the Stamp Act, which would tax domestic American commerce. Although the magnitude of these taxes was relatively trivial, American radicals argued that they reflected a sea-change in British authority and intention.

As is well-known, colonial reaction forced the British to reconsider their policy, revoking these policies in short order. In the face of the Stamp Act's failure, however, the British passed another act even as it removed this tax. The Declaratory Act of 1766 held that Parliament is unlimited and unlimitable (Greene 19**, 85), arguing that the removal of the tax was not intended

to signal that the British agreed with the Americans that they lacked the power to impose such taxes, just the opposite.

The next step in the cycle came with the passage, in 1766, of the Quartering Act. It established conditions under which colonial assemblies were to provide for British troops. When the New York Assembly (most of the troops were stationed in New York) refused to vote the money, Parliament passed the New York Suspending Act, which rendered all acts of the Assembly void until it voted out funds for troops. This Act brought forth another round of protest from the colonies.

The next major policy initiative came in 1773, with the passage of the Tea Act. To help out the East India company, Parliament lowered the duty on tea, while also beefing up the penalties for smuggling. They did so in the hope that the tax would be so low that Americans will set aside their principles and pay the duty.

Once again, the colonists protested the constitutionality of the Act. This protest culminated in the Boston Tea Party, during which the colonists destroyed tea in order to preserve constitutional claims. Once again, the British reacted to this resistance by stepping up their enforcement with passage of the Coercive or Punitive/Intolerable Acts (Greene, 23). The initial bill closed the Boston harbor until it paid for tea and until duties were collected. When this act failed to produce colonial compliance, Parliament passed three more. The Massachusetts Government Act annulled the charter of the Massachusetts colony; the Administration of Justice Act provided that colonial officials would be tried outside of the colony; and the Quartering Act allowed the imperial officials to seize property to support troops if the colonial assembly failed to allocate the necessary funds.

Finally, the Quebec Act expanded Canada south into the territory claimed by colonies, attempting to prohibit American colonial expansion into trans-Appalachia. It also legislated religious

tolerance for Catholics. Both provisions raised constitutional issues for the Americans. The prohibition on immigration unilaterally altered colonial charters by affecting their boundaries. More ominously given that many colonies were founded on sustaining religious freedom, Parliament's regulation of religion threatened the American separation of church and state. If the British had the constitutional authority to regulate religion in Quebec, it possessed the authority to do so in the American colonies.

The differing American and British conceptions of sovereignty and the constitution guided both sides in their interpretation of their own and their opponents actions. The colonists' view about the long history of divided sovereign meshed with their view of the constitution, based on precedent and long-established practice. From the colonial perspective, by breaking with precedent, the British attempts at taxation, quartering troops, suspending colonial legislatures, altering colonial charters, and regulating religious practice seemed flagrant violations of the constitution and American liberty. In the mid-1760s, the radicals predictions that the British would undermine the centrality of the colonial assemblies seemed outlandish. By the mid-1770s, the predictions had come true. The weight of the British actions thus altered the moderates' perception of British malevolence, pushing them into the radicals' camp (Rakove, Rutten, and Weingast 1999). By the end of 1774, it was hard for the moderates to sustain the view that the British were benign.

Implications of these events

From the beginning of the crisis, radicals sought radical action. But this faction was too small to be politically effective. The Tories or Loyalists occupied the other end of the political spectrum. In between these two factions stood the moderates, a group of prudent and sober individuals who

understood the risks of revolution and war (Rakove 1979, chs. 1-2). The moderates were pivotal politically: if they remained loyal, no effective challenge to the British could be mounted. But if they joined the radicals, revolution would gain sufficient support.

The events described above had a cumulative impact on moderate Americans. In our companion paper, we use a simple Bayesian model to show how the various events just noted would have altered the moderate's beliefs (Rakove, Rutten, and Weingast 1999). Whereas the moderates were inclined in the mid-1760s to believe the radicals crazy and imprudent, by the mid-1770s, many had changed their minds. Regardless of whether the British were actually intent on expropriating colonial wealth and closely regulating colonial society, the British had become a malevolent presence willing to suspend or alter the heart of the political and constitutional mechanisms that Americans believed protected their liberties and their way of life. The British willingness to regulate a range of domestic affairs — potentially including religion — and their willingness to abrogate the foundation of American liberty — the colonial assemblies and their role in creating American law — represented a direct threat to the American way of life.

The empire's division of powers served as a bright-line commitment mechanism. Understanding this principle helps explain an important component of the colonists' fears generated by the British actions and newly articulated constitutional principles after 1763. The British view that Parliament was unlimited and unlimitable threw aside the mechanisms providing credible commitment for American liberty. Moreover, from the colonial perspective, the British theory of Parliamentary sovereignty offered *no* credible substitute, not even a poorer one. This omission potentially left Americans without credible protection from opportunistic action by the metropole. Because today's parliamentary majority could not bind tomorrow's, anything enshrined in today's

legislation could be altered tomorrow. As the radicals observed early in the controversy, if the British could exercise these powers in mild and relatively innocuous ways, they could do so in ways that expropriated Americans or which began to regulate social fundamentals, such as religion.

Further, the absence of a credible compromise from the British — indeed, the absence of any attempt to court the American moderates away from the American radicals (Rakove 1979) — provided further evidence in favor of the radical's thesis. If the British were not malevolent, why did they not offer a reasonable compromise?

The role of ideas proved central to the moderates concluding that they needed to side with the radicals. The ideals of constitution, sovereignty, federalism in the empire, and above all, American liberty, gave Americans the framework within which to interpret British action. American theorists explained how these mechanisms provided credible protection under in the Empire and how the new British views directly threatened Americans. American moderates therefore came to see great danger in the British dismantling of the former mechanisms of credible commitment.

Still, a number of puzzles remain. How could the British and the Americans sustain these differences over the previous century? In particular, how could they cooperate for so long in the face of these differences? And what changed following the Seven Years' War, underpinning the explosive conflict?

3. The Self-Confirming Equilibrium Framework

This section contains the technical preliminaries that define a self-confirming equilibrium and its relation to the more traditional Nash equilibrium framework.⁸ One purpose of the approach is to allow the player's mental models of their strategic situation — as modeled by their beliefs about the extensive form — and about the intentions of others — as modeled by their strategies — to differ. These variations allow us to model players with cognitive limits in the sense identified by North (1999, ch**) — they have incorrect and inadequate understandings of their environment and about other players intentions. The equilibrium aspect of the framework provides the discipline that prevents these beliefs from being arbitrary, wildly incorrect, or mutually incompatible.

Several participants in this literature explain the motivation for this approach. Fudenberg and Levine (1993,523) weaken the Nash equilibrium requirement that each player holds correct beliefs about their opponent's strategies. They require “only that players' beliefs are correct along the equilibrium path of play. Thus each player may have incorrect beliefs about how his opponents would play in contingencies that do not arise when play follows the equilibrium, and moreover the beliefs of different players may be wrong in different ways.” In discussing these and related notions, Greenberg (1994:4) suggests that “these notions are motivated by the fact that ‘off equilibrium choices’ are not observed, and hence the requirement of ‘commonality of beliefs’ cannot be justified.”

⁸Fudenberg and Kreps (1995), Fudenberg and Levine (1993,1998), Greenberg (1994), Kalai and Lehrer (1993) explore this and related concepts. Sargent (1999) applies self-confirming equilibria to macroeconomic phenomena.

In what follows, we adapt these ideas. A principal difference is that we allow players to have differing beliefs about the extensive form, though in equilibrium, these differences occur only off the path. The idea is in part that the players have limited knowledge about one another's capabilities and they may never learn of another player's feasible move if that move is never used or observed.

The framework begins with an extensive form game, M , and a set of N players. Each player i has a set of beliefs, μ_i , about the structure of the extensive form game. These beliefs may differ from one another and from the actual extensive form, though as we will see, in equilibrium, they cannot differ "too much." Given a history of the game up to move k , h_k , a **strategy**, s_i , for player i is a rule that specifies a proposal for any move at which i may make one and is a function of h_k . Each player also has a set of beliefs, σ_{-i} , about the strategies being played by the other players on μ_i .

A **self-confirming equilibrium** is a set of strategies, $s^* = (s_1^*, \dots, s_n^*)$ and beliefs $(\mu, \sigma) = (\mu_1 \sigma_1, \dots, \mu_n \sigma_n)$ with three characteristics:

- (a) $U^i(s_i^*/\sigma_{-i}) \geq U^i(s_i/\sigma_{-i}) \quad s_i \neq s_i^* \quad \forall i$; further, this property holds for all proper subgames of μ_i ;
- (b) μ_i must be identical to M along the equilibrium path; and
- (c) $\sigma_{-i} = s_{-i}^*$ along the equilibrium path of μ_i .

Condition (a) requires that player i 's strategy maximizes her payoffs, given her beliefs about both the extensive form, μ_i , and the strategies being played by others, σ_{-i} . This condition also requires that her strategy is subgame perfect on μ_i . Conditions (b) and (c) require that no events along the path disconfirm i 's beliefs about the structure of the extensive form or about her opponents' strategies. Notice that condition (b) implies that $\mu_i = \mu_j$ along the equilibrium path, though they may differ off the equilibrium path.

Our definition of self-confirming equilibrium differs a bit from Fudenberg and Levine's, since they are interested in the asymptotic properties of learning in games. It also differs from Sargent's, since he studies a policymaker maximizing social welfare but who has an incomplete understanding (or mental model) of how the world works. It is close in spirit to Greenberg and Kalia and Lehrer.⁹

Pedagogically, one can think of a game with a self-confirming equilibrium as $n+1$ games: the game that each of the N players believes she's playing (μ_i) plus the actual game they are playing. A self-confirming equilibrium in this context can be broken down into several parts. Consider player i and the game she believes she is playing, μ_i . First, given the strategies i believes the other players are playing (σ_{-i}), i 's strategy is an optimal response. Moreover, the strategy that i believes each of the others is playing must also be an optimal response for that player on μ_i . Second, this set of strategies (s_i, σ_{-i}) constitute a subgame perfect equilibrium on μ_i . Third, the set of strategies each player actually plays (s_1^*, \dots, s_n^*) must constitute a self-confirming equilibrium on the actual game they are playing (M). That is, s_i is an optimal response for i on μ_i ; along the equilibrium path, $\mu_i = \mu_j \forall i, j$; and $\sigma_{-i} = s_{-i}^*$ along the equilibrium path of M .

To understand a bit more about self-confirming equilibria, it is useful to compare them with Nash equilibria. Self-confirming equilibria represent a generalization of Nash equilibria in the sense

⁹My approach differs from these approaches in one respect. They assume that the game tree is common knowledge. Our formulation allows for the possibility that the players have different understandings of the set of feasible moves. Of course, the self-confirming equilibrium requires that nothing about what the players observe disconfirms their beliefs, so that, in equilibrium, differences in the game tree can only occur off the equilibrium path. The idea behind this extension is twofold. First, one player may not know the entire set of feasible moves for another player; and if that player never exercises a move of which the other is unaware, then the other may never learn it was feasible. Second, this formulation allows a choice by nature of which one player is aware but another is not. Once we admit that there is no compelling reason for players beliefs about off the path play to be correct, there seems no compelling reason to assume that the game itself is common knowledge.

that all Nash equilibria are self-confirming equilibria. The latter differ from Nash equilibria in at least two respects.¹⁰ First, the Nash framework assumes that the game's structure is common knowledge. Second, this approach also assumes that the players know the other players' strategies or are aware of their ignorance. In technical terms, the common knowledge feature of Nash implies that $\forall i, \mu_i = M$. The assumption about strategies requires that, $\forall i, \sigma_{-i} = s_{-i}^*$ over all M , not just along the equilibrium path.

In contrast to the Nash framework, nothing in the self-confirming equilibrium framework forces complete agreement among the players beliefs about the structure of the extensive form and about the strategies played by others. Self-confirming equilibria allow player i to be play a strategy different from that which player j believes i is playing, where the differences occur off the equilibrium path. Because each player observes only play along the equilibrium path, neither becomes aware of their inaccurate understanding of the game. Yet these potentially contradictory beliefs are an equilibrium because no player observes anything to contradict their beliefs about μ_i or σ_{-i} . As Greenberg (1994) suggests, the players may disagree about why equilibrium arises and, further, they may not know they disagree. Player i , for example, may choose move A over move B because of her conjectures about player j 's response. But player j 's response is never observed and may in fact be different than i anticipates.

An important feature of the definition of self-confirming equilibrium constrains the beliefs about the extensive form game, μ_i , and about the strategies of others, σ_{-i} . These beliefs cannot be so wildly off that an individual is called to make a move she did not believe existed or that a player

¹⁰Fudenberg and Levine (1998, ch 6) note other differences, such as different learning properties traced through asymptotic techniques.

observes another player taking an action that the first believed infeasible (i.e., not part of μ_i). Along the equilibrium path, nothing about the play disconfirms player i 's beliefs about the extensive form or about the strategies played by the other players.

To see the effect of self-confirming equilibrium on the play of a game, consider a repeated game of the following form. Two players play a prisoners' dilemma for the first t^*-1 periods, and then in time t^* , they switch to another game with the same moves but with different payoffs so that there are no gains from cooperation. Now consider four different settings in which this game is played.

(i) Complete and perfect information: The players know there are no gains from cooperation after time t^* . This implies that, for the prisoners' dilemma portion of the game in the first t^*-1 periods, cooperation unravels in the standard way from the end. Thus, cooperation cannot be supported in the first t^*-1 periods.

(ii) Uncertainty about the whether the games switch at time t^* . Next, suppose that at time t^* the game does not automatically switch from the prisoners' dilemma to the other game, but does so with probability p . Further suppose this is common knowledge. Then, given a discount factor that is not too low and a probability of switching that is not too high, cooperation can be sustained in the early portion. Of course, if the probability of switching is large enough, cooperation cannot be sustained prior to period t^* .

(iii) Private information. Next, suppose as in (B) there is uncertainty about the switch, but there is private information about whether the switch has occurred. Specifically, suppose that both players know that one player will observe whether the switch has occurred. Then, if p is small enough, cooperation can be sustained in the first t^*-1 periods. Further, adjustment

to the new circumstance after the switch may be lagged. Suppose that p is low enough to sustain cooperation before t^* . In period t^* , the informed player knows nature's choice and will cooperate if the switch has not occurred, defect otherwise. Although the uninformed player does not know whether the switch has occurred, she will nonetheless cooperate in period t^* because p is low enough. The informed player's choice in period t^* sends a signal to the uninformed player about whether the switch has occurred. If the informed player defects in period t^* , the uninformed player knows the switch has occurred, and will defect from period t^*+1 onward. In sum, if the switch occurs, we get three phases of play: cooperation in periods 1 through t^*-1 ; the uninformed player plays cooperate in period t^* while the informed player defects; and both players defect thereafter.

(iv) Self-confirming equilibrium framework. Next suppose that the switch at time t^* is *not* common knowledge and that one of the players does not know that the switch is even a possibility. In contrast, the other player is informed about the switch. Thus, the uninformed player believes the game is just an infinitely repeated prisoners' dilemma, while the other player knows the game switches at time t^* with probability p . Assuming the discount factor is sufficiently large, then cooperation can be sustained in the early period regardless of p . The reason is that the uninformed player will cooperate because she thinks she is playing an infinitely repeated prisoners' dilemma. Given this behavior, the informed player will cooperate as well in periods 1 through t^*-1 . If nature choose to switch the games, then, as in case (C), the players will adjust to the switch asymmetrically, with the informed player defecting in time t^* and thereafter, but the uninformed player cooperating in time t^* , but defects thereafter. A final difference between this case and all the others is the "surprise" or misperception: the uninformed player is surprised by the switch in the game; and, moreover, will be initially puzzled by the other player's defection.

A model of surprises

This framework allows surprises as follows. Players in a self-confirming equilibrium may have inaccurate beliefs about the extensive form, μ_i , or the strategies being played by their opponents, σ_{-i} . As long as the system is stable, these inaccurate beliefs may not be revealed.

If the system undergoes an exogenous shock, however, then these inaccurate beliefs may come to the fore. We model this shock by allowing nature to transform the game some time after the initial self-confirming equilibrium has been in force — presumably, long after. An example is a shock that alters the payoffs for some or all of the players. The change in payoffs may then cause one or more players to alter their behavior; that is, to change their action choice in a particular move of the game. In terms of the model, something that was previously off the path becomes on the path.

Exogenous shocks of this type imply that, if one or more of the other players have inaccurate beliefs one or more of the other player's off the path play, they may mis-predict how that player reacts to the shock. Because the self-confirming equilibrium framework allows players to be mis-informed about one another's types, mis-predictions after an exogenous shock can lead to surprising behavior. As in the Nash framework, players in the self-confirming equilibrium framework may be ignorant of another player's preferences. In contrast to the Nash equilibrium framework, however, players in the self-confirming equilibrium framework may not be aware of their ignorance.

4. A Self-Confirming Equilibrium Model of the American Revolution

In this section, we apply the self-confirming equilibrium framework to the American Revolution. This application constitutes an “analytic narrative” (Bates, et al., 1998); that is, it uses a model to interpret a single case.

We begin by describing a two person game facing the British and the Americans. The basic interaction takes place in what we call the “British empire stage game.” The stage game is embedded in a context where nature has two moves, one at the beginning of the game, and one prior to the play of the stage game in a distinguished time period, t^* .

Prior to the first play of the game, Nature has the first move in which it must choose a type for the British. Although the British know that Nature makes this choice and learn its outcome, the Americans are unaware that this choice is being made. We represent this choice as a move by Nature (N_1) in figure 1. Nature’s choice of a British type results in the constitutional view held by the British, which may be either of two types. The first type is that the British retain throughout the entire period the status quo view of the constitutional that prevailed following the Glorious Revolution in the late seventeenth century and early eighteenth century (labeled, 17c). The second type is that the British constitutional view slowly evolve between the beginning of the game and period t^* into a newer constitutional view that rose to maturity in the nineteenth century (labeled 19c). The consequence of these views are described below; as suggested above, they have different implications for the understanding of British powers and actions.

The probability that Nature chooses the 17c view is p_{1j} , where $j = A, B$ for the Americans and the British, respectively. The British accurately assess the likelihood of Nature’s choice, so that $p_{1B} = p_1$, where $0 < p_1 < 1$. The Americans, however, are unaware that this choice is being made. They therefore assume that the probability is identically equal to zero; that is, $p_{1A} \equiv 0$. This first move by

Nature represents the first of two principal departures of this game from the common knowledge structure of Nash equilibrium games. By assumption, the Americans are unaware that the British type is uncertain and hence unaware that the British payoffs are uncertain. Put simply, their mental model, as expressed in their understanding of the extensive form and hence of the range of possible types they may face, inaccurately reflects their strategic environment.

Turning to the empire stage game, the British move first (see figure 2a). They must decide whether or not to intervene in the domestic affairs of the American colonies. As we will see below, there may be multiple interpretations of the meaning of this intervention, but for now we suppress them. The Americans move next, deciding whether to remain in the empire or whether to declare independence.

The stage game illustrated in figure 2a is played in every period except a distinguished period, t^* . In that period, Nature again moves, this time just prior to the play the stage game (figure 2b). Nature's (N_2) choice in this period is common knowledge and concerns the French. With a large probability p_2 , N_2 chooses that the French threat continues. But with probability $(1 - p_2)$, N_2 chooses that the French threat disappears. This choice is designed to represent the fact that the French and British were continually fighting in Europe, in North America, and elsewhere in the world for much of the 125 years between the Glorious Revolution in 1689 and the defeat of Napoleon in 1815. The French threat concerns the American colonies, and required that the French have colonies of their own in North America. The French had been a problem for the American colonies during the three-quarters of a century prior to the Seven Years' War.

The disappearance of the French threat has dramatic consequences for the interests of both the Americans and the British. Once N_2 makes its choice in period t^* , the players play the identical stage game.

Figure 3 summarizes the sequence of moves. Nature (N_1) moves first at time period t_0 . The players next play the empire stage game for the next t^*-1 periods. Then, in period t^* , they play the augmented stage game, which involves N_2 's choice of whether the French threat remains or disappears. For the remainder of game (periods t^*+1 and beyond), the Americans and British continue to play the stage game.

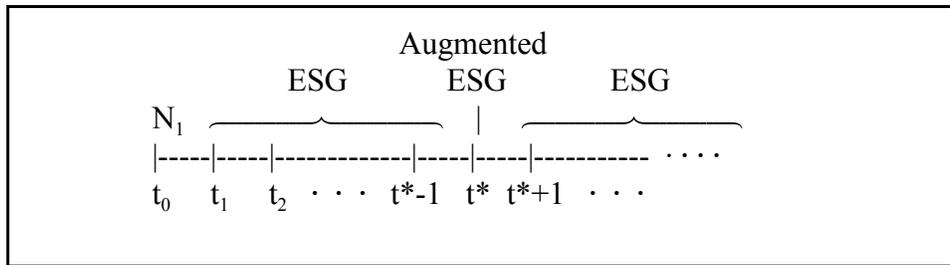


Figure 3. Time sequence of moves.

Payoffs. We now turn to the payoffs of the empire stage game. These are given in figure 4, which has four panels. Cases 1A and 1B reflect N_1 's choice that the British constitutional type is 17c; whereas cases 2A and 2B reflect N_2 's choice that the British type is 19c. For each of the two cases, panel A represents the payoffs in the presence of the French threat; while panel B represents the payoffs in the absence of the French threat.

As can be seen, the empire stage game is represented by a prisoners' dilemma in three of the four cases (all but case 2B). In all four cases, the British have the choice of intervening or not in the Americans' domestic affairs, and the Americans have the choice of remaining in the empire or declaring independence. In the presence of the French threat, both sides are better off cooperating

than non-cooperating, receiving a payoff of 5 each play (0 if they do not cooperate). Still, each player can take advantage of the other in the short-run. If the British choose unilateral defection — that is, to intervene while the Americans cooperate — the British get 7 while the Americans get -3. If the Americans unilaterally defect — that is, to declare independence — and the British cooperate, the Americans get 7 while the British receive -3. Because the British hold the 17th century view of the constitution in this case, they, like the Americans, view intervening in the domestic affairs of the American colonies as a violation of the constitution. Because N_2 chooses to maintain the French threat in case 1A, the payoffs to the stage game are constant in all periods (before and after period t^*).

In case 1B, the payoffs are the same as case 1A for time periods 1 through t^*-1 . In period t^* , N_2 chooses that the French threat disappears, so that the payoffs switch to those listed under case 1B in figure 4. The stage game under case 1B also has the structure of a prisoners' dilemma, yet the payoffs have changed. In particular, the value of cooperation between the British and the Americans has declined (though is still positive). In this case, cooperation is worth 2 to each, non-cooperation is worth .5. Put simply, in the absence of the French threat, the Americans and the British need each other less. Unilateral defection gives the defector 2.5, while the sucker a payoff of -1.

In case 2A, the payoffs are identical to those in case 1A. Although N_1 chooses the British type as 19c, this does not affect the game, as N_2 chooses to retain the presence of the French threat.

The main difference in the in the cases occurs in case 2B, when N_1 chooses the British constitutional type as 19c and N_2 chooses to remove the French threat. The strategic situation here differs considerably from the other three cases. As in case 1B, the move labeled cooperation is worth 2 to both players. Unilateral defection by the Americans has the same payoffs as under case 1B, 2.5

to the Americans and - 1 to the British. The principal changes in payoffs occur when the British defect. Unilateral British defection represents intervention into the American economy. When the British type is 19c, its payoff to this defection is 6, not 2.5 as in case 1B. If both players defect, the Americans receive .5 and the British, 3.

Equilibria. This game possesses multiple equilibria, and we do not attempt a full characterization. Mutual defection, for example, is an equilibrium throughout. We are interested in the ability of the players to maintain cooperation, particularly as their environment changes as represented by nature's choice about the French threat. Hence we focus on equilibria in which cooperation can be sustained.

Following the definition of self-confirming equilibrium, we write the strategies contingent on those choices by nature of which the players are aware. Thus, the British strategy will depend on both choices by Nature. In contrast, the Americans are unaware of Nature's first choice, so their strategy will not be contingent on N_1 's choice. Thus, the game the Americans believe they are playing differs from the game the British believe they are playing; i.e., $\mu_A \neq \mu_B$.

Our strategy in solving this game is to suggest that the Americans and the British were in a self-confirming equilibrium for generations prior to the distinguished time period, t^* , when nature removed the French threat. Because the Americans are unaware that Nature chose a British type, they do not accurately predict British behavior under all circumstances following Nature's choice about the French threat at time t^* . To be a self-confirming equilibrium, however, nothing about the British behavior disconfirms the Americans' beliefs about the British strategy or the structure of the game.

The equilibria are described as follows. First, suppose that N_1 chooses 17c. In this case, for δ sufficiently large, cooperation can be sustained, regardless of whether N_2 chooses to continue or

remove the French threat. Cooperation is supported by the usual trigger strategies so that each player cooperates in period 1 and, thereafter, cooperates as long as their partner does so in the previous period. If ever a player's partner fails to cooperate, then that player will defect forever. Further, players who fail to punish when called on to do so are themselves punished.

Second, suppose that N_2 chooses 19c. In this case, the British strategy is contingent on N_2 's choice in a different way. In the absence of the French threat, the value of intervening in American domestic affairs is larger for the British. And given the low payoff from cooperation relative to British defection, cooperation cannot be sustained. Thus, contingent on N_2 's choice to remove the French threat (ϕ), the British switch their actions from non-intervention to intervention.

The Americans, on the other hand, are blithely unaware in this case that the British constitutional view is 19c, not 17c. According to their beliefs, μ_A , they are in case 1B. The American's strategy is therefore the same as in case 1B; and further, they believe that the British strategy is the same as in case 1B. The Americans believe that the British have an interest in cooperation in the absence of the French threat. Whereas in fact, when the British are type 19c, they do not want to cooperate. Thus, a real surprise emerges in Case 2B at time t^* , when N_2 chooses to remove the French threat and the British choose to intervene in the American domestic affairs.

This equilibrium encompasses a pair of mutually inconsistent expectations in case 2B where the French threat disappears. The British believe they have the power to intervene and have no reason to believe that the Americans will resist. Their conjecture is wrong, but for generations prior to the removal of the French threat, this part of the game tree was off the path and thus never observed. The Americans were unaware of the uncertainty about the British type and thus that the British constitutional beliefs had changed. As the British never chose to exercise the authority they

believed they had, the Americans never obtained evidence of the new British constitutional views. In short, both sides mis-conjecture in equilibrium about the other's behavior.

Implications for the American Revolution

Our application of the model to the case is as follows. Nature chose 19c as the constitutional type for the British. Although the Americans remained ignorant of this, the British constitutional view evolved from that in the 17th century at the beginning of the game to the beginnings of the 19th century view by period t^* , representing the defeat of the French at the end of the Seven Years' War in 1763. For a century, the Americans and the British were in a self-confirming equilibrium equilibria where the Americans misunderstood their strategic situation. As long as the French threat remained, however, their misunderstanding did not matter. Both the Americans and the British had a sufficiently strong interest in maintaining cooperation. Thus, by the end of this 100 years, the British had evolved the new 19c constitutional view and thus believed they had the power to intervene in the American domestic affairs. As long as the French threat remained, however, they chose not to do so.

The Seven Years' War fundamentally altered British interests in several ways. The demise of the French lowered both side's willingness to bear costs to maintain their relationship. Simply put, they needed each other less. Moreover, after the war, the British had a multifaceted empire, no longer dominated by the Americans. The British had to manage French Canada and the Indians in trans-Appalachia. Rules designed by the British to manage the whole might not be optimal for the participants in any of the parts. Finally, the British ended the war with a huge debt and sought to have the Americans pay what they believed was a modest share of this cost.

Following the demise of the French threat, therefore, the British switched choices, seeking to intervene in the colonial domestic economy. Because the Americans did not know the British type, the Americans interpret this action as a violation of the constitution and thus as a defection from mutual cooperation. The Americans then exercised the contingent trigger in their equilibrium strategy and declare independence. The British, however, did not believe that they had defected from cooperation. Moreover, because they had no reason to believe that the Americans misunderstood this, they believed that the Americans would pay the small tax to finance their defense.

The British therefore believed, in (self-confirming) equilibrium, that the Americans would acquiesce in their intervention. When instead the Americans resisted, the British interpret the Americans' resistance as defection from the empire. They reacted, in turn, with harsher measures designed to punish and deter American resistance. The British mis-predicted the American reaction however. Instead of deterring American resistance, British actions strengthened it. The consequence is that both players switched to mutual punishment, and war ensued.

The key to the model is that there are two patterns of behavior, depending on the combination of nature's choices. When N_1 chooses the new type (19c) for the British, the Americans are not aware of it. As long as the French threat remains, however, the British never act on the changes inherent in their type. In this self-confirming equilibrium world, the Americans and the British held contradictory theories about the world. The two sides' beliefs about the extensive form, μ_A and μ_B , differed from one another. Further, both sides have inaccurate beliefs about the other's strategy after time t^* . The Americans' mistakenly believe that the British strategy have an interest in continuing to cooperate despite the absence of the French threat. The British mistakenly believe that the

Americans will accept their intervention in the form of modest taxes. And yet, as long as the French threat remained, neither side acted in a way that disconfirmed the other's beliefs.

Only when the French threat disappears does conflict emerge and thus the two side's different constitutional interpretations become relevant. Because the Americans were unaware of the changes in the British constitutional view, they did not correctly anticipate British behavior following the removal of the French threat. The British view and action came as a surprise — indeed, a rude shock — forcing American to radical action.

One may ask whether the above behavior is really an equilibrium — after all, both the Americans and the British are “surprised” in the play after period t^* , the defeat of the French. Clearly both sides learn at this time that something in their world view is false. In a technical sense, the strategies defined before period t^* are not an equilibrium since they cannot survive the data observed after t^* . In a practical sense, however, we believe that this application makes sense. The period t^* takes place over 100 years after the American colonies were established and over 75 years after the British and the French began fighting in 1689. This represents three generations of mutual coexistence in which nothing either side did disconfirmed the others' beliefs. The behavior modeled by the empire stage game in case 2A prior time t^* thus occurred over several generations. As our model suggests, this behavior is stable, and reflects the self-confirming equilibrium feature: The British type (and hence payoffs) differed from that anticipated by the Americans, although these differences come into play only once Nature removes the French threat. Nothing the British did during the generations prior to time t^* disconfirmed the Americans' beliefs about the British. Although the British believe they have the authority to intervene in the domestic affairs of the American colonies, they chose not do so long as the French remained a threat.

Notice that the different constitutional views imply different interpretation of the British and American moves. To a 17c British type, intervention in the American colonial economy represents expropriation and opportunistic behavior. Doing so would thus alter constitution precedent throughout the empire, including in Britain itself, an outcome they would want to avoid. To a 19c British type, however, intervention represents a discretionary policy choice without constitutional implications elsewhere in the empire or Britain. In this case, intervention has no negative normative content for the British.

Nature's choice of a British type also affects the British interpretation of the American behavior. Because a 19c British type views intervention as a matter of policy, the American resistance to its intervention is treason. To the Americans, the British violation of their interpretation of the constitution is tyrannous.

The self-confirming equilibrium model thus helps explain how the two sides could cooperate for several generations, only to suddenly explode into mutual antagonism, resistance and punishment, and then revolutionary war. Moreover, this view helps explain the connection between the realm of ideas emphasized by the historians and the realm of action. The American theories of the empire's constitution reflect their attempts to understand how the British credibly committed to protect American rights in the previous era. According to this theory, British intervention — however modest — was precedent-setting and thus undid the institutions protecting American liberty and rights. The British articulation of their new constitutional view exacerbated the American fears. The notion that Parliament was unlimited in its powers over the empire implied that it new no institutional bounds on its possible intervention in America. Put simply, this said to the Americans that no credible commitments to American rights existed or were possible.

The self-confirming equilibrium view also helps explain a critical aspect of the controversy, namely that both sides believed themselves completely in the right and the other at fault. The Americans, given their constitutional view and understanding about credible commitments, saw the British intervention and new constitutional view as threatening their ways of life. They therefore sought to defend themselves. The British — facing the problems of governing a new, huge, and complex empire -- sought to alter a few rules governing the Americans. They were surprised at the vehemence of the American reaction and in short order came to interpret American resistance as a move toward independence — that is, defection from the empire. They reacted harshly, helping to convince moderate Americans in the radicals' views that revolutionary action was the best path (Rakove, Rutten, and Weingast 1999).

5. Conclusions

In this paper, we utilize the concept of self-confirming equilibrium to model a particular type of cognitive limit on rational action — the notion that individuals' mental models of their strategic environment often remains fundamentally incomplete and inaccurate. The approach requires that the players' misunderstandings cannot be arbitrary, but must instead with stand the test of time; that is, the players' misunderstandings must be compatible with the evidence resulting from their interaction along the path of the game.

Self-confirming equilibrium thus restricts players' inaccurate beliefs about one another's behavior to be about choices off the equilibrium path. Rational action in this perspective is often

based on inaccurate counterfactuals: an individual chooses A over B because she prefers the consequences of choosing A over the consequences of B. This implies, however, that the individual never observes the sequence of action in the game implied by B. If the first player's predictions about the other's reactions under B are inaccurate, the individual may never learn of this.

Surprises can emerge in this framework when an exogenous shock affects the game. For example, if the shock alters preferences, one player may choose a move that was once off the path and thus unexpected by at least one of the other players.

The self-confirming equilibrium approach adds to our understanding of the American Revolution in several ways. First, our perspective provides a link between the ideals discussed by historians with everyday politics and economics through the concept of credible commitment. The Americans interpreted the British behavior after the Seven Years' War as altering the fundamental institutions that for the previous century, institutions that had provided for credible commitment against arbitrary authority, expropriation, and opportunism. American liberty centered around the colonial assemblies. Whether Americans most valued their religious freedom, overseas trade, or the protection of property rights in land or slaves, the colonial assembly was critical to protecting them. British parliamentary sovereignty threatened this system in principle, and the new British policies governing the empire threatened it in practice.

At the beginning of the crisis in the early and mid-1760s, the radicals' claims that the British would destroy American liberties, including their legislatures, seemed outlandish. By the mid-1770s, many of the radical's dire predictions had come true. The British demonstrated a clear willingness to regulate a large variety of American domestic policy, and to punish American resistance by disbanding the colonial assemblies.

Second, the perspective helps clarify the history of the American Revolution, demonstrated the critical link between the realm of ideas, as emphasized by historians, and the realm of action, as occurred in the revolution. The constitutional ideals represented different understandings of the credible commitments protecting American rights and liberty. In the American view, the British behavior was tyrannous and pointed toward exploitation. As Rakove, Rutten, and Weingast (1999) emphasize, British attempts to punish American resistance — particularly the British suspension of American colonial assemblies — provided stark evidence of this perspective.

Our model also points to how the Seven Years' War fundamentally altered each side's interests, lowering the value of cooperation to each. Although historians recognize the importance of the shift at the end of the Seven Years' War, with the exception of Tucker and Hendrickson (1982), they do not analyze its full impact on the revolutionary crisis. In particular, they do not investigate how the demise of the French threat altered each sides' interests, thus altering the basis for cooperation within the empire. Our perspective suggests how the war fundamentally altered both sides' interest in the empire and how this change was critical for war.

Third and perhaps most importantly, our model explains how the two sides could cooperate peacefully for over 100 years, only then explode in war and revolution. The two sides had long held mutually inconsistent idea and ideals; yet these differences did not matter for the century prior to the Seven Years' War. Although the British believed they had the authority to alter the governance

structures of the empire, they had no interest in doing so.¹¹ Only after the war did they seek to restructure the empire, with dire consequences.

Historians emphasize the evolution of the American and British ideas and how they clashed (Greene 1986, Woods 1969). They do not, however, provide the mechanism that explains how differences in ideals, even big ones, led to revolution. Our approach shows how the Americans felt threatened by the British actions while the British viewed their policies as legitimate.

Fourth, beyond the differences in ideals, our perspective emphasizes the importance of the lack of common knowledge across the Atlantic: the differences in ideals emerged over a 100 year period in a way that neither perceived the other's perspective. The self-confirming component of our analysis is critical. The self-confirming equilibrium framework thus helps explain an aspect of the revolution ignored by historians, how the two sides interact for several generations and yet now understand one another. During their peaceful cooperation in the presence of the French threat, neither side generated any evidence of the contradictions inherent in their differing perspectives.

The perspective elucidates what made the Americans so fearful and angry about the British behavior — the British constitutional interpretation combined with their behavior to imply an absence of credible commitments to protecting American values and liberty (Rakove, Rutten and Weingast 1999). The British willingness to attack the colonial assemblies fundamentally altered the American domestic environment. Most rights and liberties of Americans centered around their assemblies. The British seeming arbitrary willingness to dismantle these assemblies — without any

¹¹Recent scholarship emphasizes that the changes in the structure of the empire after the Seven Years' War can be traced earlier to events beginning before the war (see, for example, ***; and also Tucker and Hendrickson 1984). But these events do not present anything like the profound challenge to the American ideas as the explicit British intervention following the war.

substitute doctrine to provide a credible commitment to American rights — implied that literally everything became at stake in the colonies. As we show in our companion paper, the high stakes helped sway the moderates into the radicals' camp (Rakove, Rutten, and Weingast 1999).

Our view also adds to the historians approach by explaining how the British and the Americans could have very different interpretations of the same events during the revolutionary crisis. To the Americans — interpreting and acting under their views drawing on the seventeenth century British constitution — long-established practice and precedent attained constitutional status for their sovereignty over domestic affairs. The federal structure of the empire provided the basis for credible commitment to a variety of American rights and liberties, including religious freedom, freedom of contract, and rights in land and slaves. In this view, British behavior after the Seven Years' War violated the constitution and directly threatened Americans' daily lives. In this view, British behavior was therefore tyranny.

To the British, interpreting and acting under their views of what was to become the mature nineteenth century view of the constitution, the question of intervention in the American colonies was a matter of policy choice. For decades they had chosen not to intervene. Yet the absence of exercising their authority to intervene, in the British view, in no way limited their future ability to do so, as they so clearly stated in the Declaratory Act of 1766, which at once removed the stamp tax but declared that Parliament had authority to impose such taxes. Following the Seven Years' War, the world had changed, leaving the British with unprecedented debts, a huge new empire to govern, and the belief that the Americans should help pay. The choice facing them was one of a modest intervention or not. From the British perspective, the American resistance to moderate and sensible British policy was treason.

The self-confirming equilibrium perspective puts the onus in part on the player's inaccurate mental models, their inability to have the complete and accurate of their strategic situation. An outside observer who could see both sides, as we do now, can understand how these differences emerged, were sustained, and then came into conflict. But neither side had this "birds eye view" perspective (Hutchins 1995), instead interpreting the other's actions and ideas as defecting from cooperation.

The self-confirming equilibrium perspective does not allow just any misunderstanding to survive, however. Because many inaccurate mental models lead immediately to surprises, these events force individuals to update their views. By contrast, in the American Revolution, many ideas had evolved from practices over 100 years, implying that the contradictions in ideas and interpretations had to have self-confirming equilibrium's non-disconfirming property. The Americans and the British failed to see that they were developing very different constitutional understandings of the empire, in part because it had never mattered.

The advantage of the self-confirming equilibrium approach is twofold. First, it provides a specific tool for analyzing the interaction of rational self-interested players when their mental models of the world and of one another are incomplete and inaccurate. The equilibrium aspect of the approach imposes discipline on the types of mental models by requiring that they meet the test of time, that individuals mental models must be consistent with the evidence they receive, even if the models are imperfect.

Second, the self-confirming equilibrium approach gives us a tool for analyzing how each side in a dispute can believe they are acting reasonably while the other is not. Although the players may

have differing mental models — and hence different ideas and interpretations of their worlds — they may never realize this until it is too late. As long as the environment remained stable, the discrepancy never became apparent. When the world changes in some fundamental way, disturbing a previously stable self-confirming equilibrium, the discrepancy in mental models becomes apparent. In the face of an environmental shock, one of the players may choose an action unanticipated by the other. This can cause significant mutual misunderstandings, allowing disputes to escalate.

In the case of the American Revolution, both sides believed they were behaving honorably and that the other side unreasonably. Each believed they were cooperating while the other was defecting. In the Nash equilibrium framework, no such ambiguity exists: defection is either obvious or hidden, but if the latter, every player knows the extent to which it is hidden. Both sides escalated in reaction to the other's apparent and persistent unreasonableness. And both sides viewed the other at fault. From our bird's-eye-view perspective, the problem looks symmetric; but at the time neither side of the Atlantic saw it that way. The self-confirming equilibrium framework shows how players can have these inconsistent beliefs about one another.

References

- Bailyn, Bernard. 1967. *Ideological Origins of the American Revolution*. Cambridge: Harvard University Press.
- Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry R. Weingast. 1998. *Analytic Narratives*. Princeton: Princeton University Press.
- Beard, Charles A. 1913. *An Economic Interpretation of the Constitution of the United States*.
- Blackstone, William. 1765. *Commentaries on the Laws of England. A Facsimile of the First Edition of 1765-1769. Vol. 1: Of the Rights of Persons*. Chicago: University of Chicago Press.
- Christie, I.R. 1966. *Crisis of Empire: Great Britain and the American Colonies, 1754-1783*. New York: W.W. Norton.
- Clark, Andy. 1997a. "Economic Reason: The Interplay of Individual Learning and External Structure," in John N. Drobak and John V.C. Nye, eds., *The Frontiers of the New Institutional Economics*. San Diego: Academic Press.
- Clark, Andy. 1997b. *Being There: Putting Brain, Body, and World Together Again*. Cambridge: M.I.T. Press.
- Cook, Karen, and Margaret Levi. 1990. *The Limits of Rationality*. Chicago : University of Chicago Press.
- Denzau, Arthur T., and Douglass C. North. 1994. "Shared Mental Models: Ideologies and Institutions," *Kyklos* 47: 3-31.
- Elster, Jon. 1998. "Emotions and Economic Theory" *J. of Economic Literature*. XXXVI (March) pp 47-74.
- Frank, Robert. 19** *Choosing the Right Pond*
- Frank, Robert. 19**. "Emotions" *AER*
- Frey, Bruno. 1997. *Not Just for the Money: An Economic Theory of Personal Motivation*. Cheltenham, UK: Edward Elgar.
- Fudenberg, Drew, and David M. Kreps. 1988. "Learning, Experimentation and Equilibrium in Games." Mimeo, Stanford University.
- Fudenberg, D., and D. Kreps (1995) "Learning in Extensive-Form Games I. Self-confirming equilibrium" *Games and Economic Behavior*. 8: 20-55.
- Fudenberg, Drew, and David K. Levine. 1993. "Self-Confirming Equilibrium" *Econometrica* 54: 1375-85.
- Fudenberg, Drew, and David K. Levine. 1998. *The Theory of Learning in Games* Cambridge: M.I.T. Press.

- Greenberg, J. 1994. "Social Situations without Commonality of Beliefs: Worlds Apart but Acting Together." Working Paper, McGill University.
- Greene, Jack P. 1986. *Peripheries and Center: Constitutional Development in the Extended Politics of the British Empire and the United States, 1607-1788*. New York: W.W. Norton.
- Greene, Jack. 1991. *Negotiated Authorities: Essays in Colonial Political and Constitutional History*. Charlottesville: University of Virginia Press.
- Greene, Jack. 1995. *Interpreting Early America: Historiographical Essays*. Charlottesville: University of Virginia Press.
- Hayek, Frederick. 1960. *The Constitution of Liberty*. Chicago: University of Chicago Press.
- Hutchins, Edwin. 1995. *Cognition in the Wild*. Cambridge: M.I.T. Press.
- Kalai, Ehud, and Ehud Lehrer. 1993. "Subjective Equilibrium in Repeated Games," *Econometrica*.(September) 61: 1231-40.
- Kannemen, Daniel, and Amos Tversky. 1979. "Prospect Theory." *Econometrica*.
- Lupia, Arthur, and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* New York: Cambridge University Press.
- Mansbridge, Jane J. 1990. *Beyond Self-Interest*. Chicago: University of Chicago Press
- McCusker, John J., and Russell R. Menard. 1985. *The Economy of British America, 1607-1789*. Chapel Hill: University of North Carolina Press.
- Milgrom and Roberts (19**)
- Morgan, Edmund S. 1956. *The Birth of the Republic, 1763-89*. Chicago: University of Chicago Press.
- North, Douglass C. 1999a. "The Process of Economic Change." Mimeo., Washington University.
- North, Douglass C. 1999b. "Hayek's Contribution to Understanding the Process of Economic Change," Working Paper, Hoover Institution.
- North, Douglas C., William Summerhill, and Barry R. Weingast. 1999. "Order, Disorder and Economic Change: Latin America vs. North America," in Bruce Bueno de Mesquita and Hilton Root, eds., *Governing for Prosperity*. Yale University Press (forthcoming).
- Potter, Janice. 1983. *The Liberty We Seek: Loyalist Ideology in Colonial New York and Massachusetts*. Cambridge: Harvard University Press.
- Rakove, Jack N. 1979. *The Beginnings of National Politics: An Interpretive History of the Continental Congress*. Baltimore: Johns Hopkins University Press.
- Rakove, Jack N. 1996. *Original Meanings: Politics and Ideas in the Making of the Constitution*. New York: Knopf.

- Rakove, Jack, Andrew Ruten, and Barry R. Weingast. 1999. "Ideas, Interests, and Credible Commitments in the American Revolution" Working Paper, Hoover Institution, Stanford University.
- Reid, John Phillip. 1995. *Constitutional History of the American Revolution*. Madison: University of Wisconsin Press.
- Sargent, Thomas J. 1993. *Bounded Rationality in Macroeconomics*. New York: Oxford University Press.
- Sargent, Thomas J. 1999. *The Conquest of American Inflation*. Princeton: Princeton University Press.
- Satz, D., and J. Ferejohn. "Rational Choice and Social Theory" *J. of Philosophy*. (1994) 9102: 71-87.
- Smith, Joseph Henry. 1965. *Appeals to the Privy Council from the American Plantations*. (New York: Octagon Books).
- Tucker, Robert W., and David C. Hendrickson, 1982. *The Fall of the First British Empire: Origins of the War of American Independence*. Baltimore: Johns Hopkins University Press.
- Weingast, Barry R. 1995. "The Economic Role of Political Institutions." *J. of Law, Economics, and Organization*
- Wood, Gordon S. 1966. "Rhetoric and Reality in the American Revolution," *William and Mary Quarterly* 3rd. ser. 23: 3.
- Wood, Gordon S. 1969. *The Creation of the American Republic, 1776-1787*. New York: W.W. Norton.

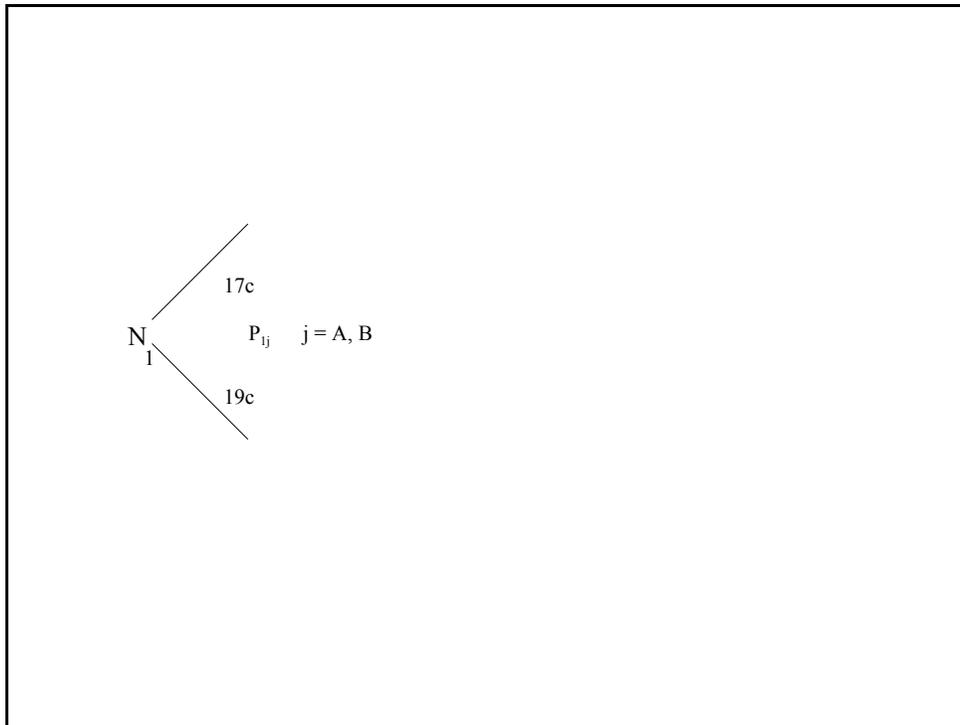


Figure 1. Nature's Choice of a British Type.

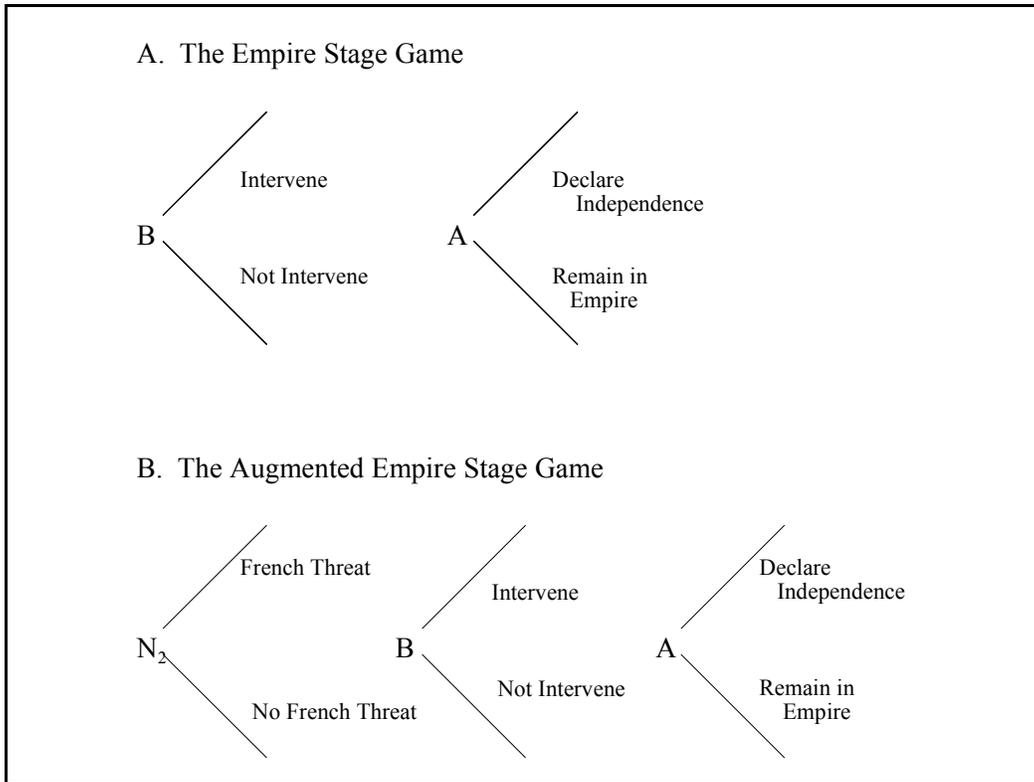


Figure 2. The Empire Stage Game.

Figure 4. Payoffs from the Empire Stage Game

The two (independent) choices by nature determine four separate cases, whether the British constitutional type is that of the seventeenth century (17c) or the nineteenth century (19c) and whether there is a french threat (Fr Th) or not (φ).

Case 1A: 17c, and French Threat

		British	
		Not Intervene	Intervene
Americans	Remain	5,5	-3,7
	Declare Independence	7,-3	0,0

Case 1B: 17c, and No French Threat

		British	
		Not Intervene	Intervene
Americans	Remain	2,2	-1,2.5
	Declare Independence	2.5,-1	.5,.5

Case 2A: 17c, and French Threat

		British	
		Not Intervene	Intervene
Americans	Remain	5,5	-3,7
	Declare Independence	7,-3	0,0

Case 2B: 17c, and French Threat

		British	
		Not Intervene	Intervene
Americans	Remain	2,2	-1,6
	Declare Independence	2.5,-1	.5,3