

NBER WORKING PAPER SERIES

CONSTITUTION OR CONFLICT?

Herschel I. Grossman

Working Paper 8733
<http://www.nber.org/papers/w8733>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2002

I have received helpful comments from Gerard Alexander, Murat Iyigun, Martin Kolmar, and Juan Mendoza, from my colleagues in the Brown Political Economy Workshop, and from participants in seminars at the University of Maryland, Korea University, Korea Development Institute, and Seoul National University. The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

© 2002 by Herschel I. Grossman. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Constitution or Conflict?
Herschel I. Grossman
NBER Working Paper No. 8733
January 2002
JEL No. D7

ABSTRACT

A self-enforcing constitution creates a political process that provides an alternative to civil conflict for resolving disputes among the constituent groups of the polity. This paper is concerned with discovering the conditions under which it is possible to design such a self-enforcing constitution. The paper is also concerned with discovering generic features of a self-enforcing constitution. The analysis yields the following theoretical propositions: If and only if (1) none of the parties to a dispute regards the dispute to be too important relative to the expected incremental cost of civil conflict and (2) no party has too big of an advantage in civil conflict, then the parties are able to resolve a dispute constitutionally. Also, under a constitution that is self-enforcing the outcomes of constitutional contests for political power do not matter too much. The paper illustrates the relevance of the theoretical analysis by applying these propositions to two dramatic historical examples of constitutional failure: the secession of eleven Southern states from the Union in 1861 and the National Socialist revolution in Germany in 1933.

Herschel I. Grossman
Department of Economics
Box B
Brown University
Providence, RI 02912
and NBER
herschel_grossman@brown.edu

A constitution attempts to create a political process for resolving disputes among the constituent groups of a polity. From the perspective of the polity as a whole constitutional resolution of disputes is desirable because it avoids the incremental costs of civil conflict. These incremental costs can include the use and the destruction of scarce resources and the loss of life.

Constitutions sometimes succeed and sometimes fail. The problem is that the political process prescribed by a constitution provides a viable alternative to civil conflict if and only if the parties to the dispute to be resolved voluntarily choose to abide by the constitution — that is, if and only if the constitution is self enforcing. If any party to a dispute does not voluntarily accept the outcome of the constitutional political process, even if this party can be coerced into complying, then the parties are left to resolve the dispute by civil conflict.

This paper is concerned with discovering the conditions under which it is possible to design a constitutional political process that can resolve a given dispute without civil conflict — that is, the conditions under which in the face of given dispute the set of self-enforcing constitutions is not empty. The paper is also concerned with discovering some generic features of a self-enforcing constitution.

Overview

A constitution includes two essential components: the specification of the form and timing of constitutional contests for political power and the specification of the prerogatives of winners of constitutional contests. These components can be themselves the subject of the constitution, or they can be derived from general principles expressed in the constitution. In addition, these components, or the general principles from which they are derived, can be embodied either in explicit provisions of the constitution or in implicit understandings.

Both of the essential components of actual constitutions vary widely. In electoral democracies constitutional contests involve periodic competition for the votes of an electorate. In contrast, in aristocracies and autocracies constitutional contests involve competition for the favor of wise men, elders, or hereditary rulers. Many constitutions combine democratic

and aristocratic features. For example, under the Constitution of the United States a recent inconclusive electoral competition led to a second competition in which the candidates for president competed for the favor of a court of appointed judges. This paper does not draw a sharp distinction either between democratic and nondemocratic constitutions, a distinction that is often ambiguous, or between explicit constitutional provisions and implicit understandings.

To anticipate the main results of the theoretical analysis, I find that it is possible to design a constitutional political process that can resolve a given dispute without civil conflict if and only if no group regards the dispute to be too important relative to the expected incremental costs of civil conflict and no group has too big of an advantage in civil conflict. The analysis also shows how a constitution that provides a viable alternative to civil conflict must limit the prerogatives of winners of constitutional contests such that the outcomes of constitutional contests for political power do not matter too much.

The paper illustrates the relevance of the theoretical analysis by using these results to explicate two dramatic historical examples of constitutional failure. One example is the secession of eleven Southern states from the Union in 1861 and the ensuing American Civil War. In this example the theory helps us to identify changes in exogenous factors that, after seventy years in which artful compromises had allowed disputes over slavery to be resolved constitutionally, prevented another successful renegotiation of the constitution. The other example is the National Socialist revolution in Germany in 1933. In this example the theory suggests why, despite repeated attempts to deal with divisions among economic and social interests constitutionally, in the end there was no viable alternative to civil conflict for resolving the dispute over the demand for a Nazi dictatorship.

The main objective of this paper — to discover the conditions under which the set of self-enforcing constitutions is not empty — is logically prior to the following two related questions. First, if the set of self-enforcing constitutions is not empty, is the possibility of designing a self-enforcing constitution realized? Second, which constitution is chosen from

the set of self-enforcing constitutions, if that set includes more than one element? This paper determines the preconditions for addressing these important questions.¹

A Generic Dispute

Assume that a polity faces a recurring dispute between two constituent groups over a specific economic or social issue.² These constituent groups can be either narrow, like rival ruling elites or rival political cliques, or broad, like ethnic groups or social classes.

Possible examples of disputes include the following:

- A political squabble: the share of political patronage that goes to the politicians associated with one clique rather than another.
- A kleptocratic rivalry: the share of kleptocratic rent that goes to one ruling elite rather than another.

¹The answers to both of these questions depend on, among other things, whether a consensus of the constituent groups of a polity is necessary to establish a constitution, or whether one constituent group can impose a constitution on others, or an outsider can impose a constitution. Constitutional choice by consensus has been the subject of much recent research. For example, Avinash Dixit, Gene Grossman, and Faruk Gul (2000), who generalize the seminal work of Alberto Alesina (1988), pose as alternatives a constitution that specifies limits on the prerogatives of the party in power and a constitution without such limits, implicitly assuming that both of these constitutions would be viable. This work does not consider a constitution to be an alternative to civil conflict, as in the present paper. An extensive literature in political science also addresses the question of constitutional choice. See Gerard Alexander (2001) for an overview and recent contribution to this literature.

In work that is more closely related to the present paper Daron Acemoglu and James Robinson (2001) and Adam Przeworski (1991, 2001) ask whether a constitution can enable a polity to avoid civil conflict. But, this work takes civil conflict to be only a mechanism for switching between democratic and nondemocratic constitutions. This work does not view a constitution and civil conflict to be alternative ongoing methods for resolving disputes, as in the present paper.

²A more general model would consider a set of disputes involving many issues and many groups. In practice some disputes might be resolvable constitutionally, while other disputes lead to civil conflict. In this event a constitution might continue to function during a civil conflict.

- A divisive distributional issue: the share of national income that goes to one social class rather than another.
- A divisive ideological issue: the degree of tolerance for the religious practices of one ethnic group, these practices being offensive to another ethnic group.

Let Party A and Party B denote the groups involved in the dispute. The term “Party” is used here in the generic sense of a group with a common purpose. Party A and Party B do not necessarily correspond to political parties.

Let X , $X \in [0, 1]$, denote the realization of the disputed issue. Party A prefers X to be larger, whereas Party B prefers X to be smaller. To implement this difference in preferences, assume that in each iteration of the dispute the realization X adds the amount αX , $\alpha \in (0, \infty)$, to the utility of Party A and the amount $\beta(1 - X)$, $\beta \in (0, \infty)$, to the utility of Party B. The preference parameters, α and β , are weights that calibrate the importance of the dispute about X for each party. This formulation implicitly assumes that the choice of X in each iteration of the dispute affects only the current utility of the Parties and that the choice of X is costlessly reversible in the next iteration. Also assume that each party knows how important the dispute is for the other party as well as for itself.

A Constitution

Consider a constitution that prescribes a periodical constitutional contest to determine which Party has the constitutional prerogative to choose X until the next constitutional contest. Suppose that under this constitution Party A has probability P_A of winning each constitutional contest and that Party B has probability P_B of winning each constitutional contest, where $P_A + P_B = 1$. Assume that these probabilities, which depend on, among other things, the design of the constitution, are known to both Parties.

This constitution also limits the discretion of the winning Party in choosing X . As mentioned above, these limits can be embodied either in explicit provisions of the constitution or in implicit understandings that supplement these explicit provisions. Let these limits be

that Party A, if it wins a constitutional contest, will not set X larger than X_A , and that Party B, if it wins a constitutional contest, will not set X smaller than X_B . For the purposes of the present analysis, the quadruple, S , where $S = \{P_A, P_B, X_A, X_B\}$, completely describes the constitution.

To contrast a constitution with a civil conflict, let Q_A denote the probability that Party A would win a civil conflict with Party B, and let Q_B denote the probability that Party B would win such a civil conflict, where $Q_A + Q_B = 1$. Assume that these probabilities are exogenous and known to both Parties.³ Assume also that, because the winner of a civil conflict subjugates the loser, if Party A were to win a civil conflict, then it could set X equal to one, its most preferred value, whereas, if Party B were to win a civil conflict, then it could set X equal to zero, its most preferred value.

Finally, normalize the costs of a constitutional contest to zero, and let the positive numbers, C_A and C_B , which are calibrated in units of utility, denote the expected incremental costs of a civil conflict to the respective Parties. As mentioned above, these incremental costs can include the use and the destruction of scarce resources and the loss of life. For simplicity, the analysis takes C_A and C_B to be exogenous.⁴

Within this schema constitutional resolution of a dispute differs from a civil conflict in three respects. First, the probabilities associated with winning a constitutional contest are socially constructed. Second, a constitution limits the prerogatives of the winner of a constitutional contest. Third, a constitutional contest is less costly than a civil conflict.

³An interesting extension of the model would be to allow Q_A and Q_B to be state variables whose current values depend on past realizations of constitutional contests and civil conflicts.

⁴See Dmitriy Gershenson and Herschel Grossman (2000) and Grossman (1999) for examples of models in which the resource cost of civil conflict is endogenized. An alternative way to make the expected cost of civil conflict larger than the expected cost of a constitutional contest would be to assume that the Parties are risk averse and that under a constitution, but not with civil conflict, the Parties can implement supplementary understandings that decrease the variance of X . Alesina (1988) and Dixit, Grossman, and Gul (2000) explore the possibility of collusion by political parties to decrease risk.

Constitutional Commitments

Assume, for the moment, that both Parties were committed to abide by the constitution. Under this assumption it would be possible to design a constitution that is attractive to both Parties if and only if there exists one or more quadruples, S , such that for both Parties in any iteration of the dispute the expected contribution to utility from using a constitutional contest to resolve the dispute about X would be larger than the expected contribution to utility from a civil conflict. Given the properties of a civil conflict, such quadruples, S , must satisfy the following two conditions:

$$(1) \quad \alpha (P_A X_A + P_B X_B) > \alpha Q_A - C_A.$$

$$(2) \quad \beta [P_B(1 - X_B) + P_A(1 - X_A)] > \beta Q_B - C_B.$$

With X_A and X_B as binding constraints, the LHS of condition (1) is the expected contribution to the utility of Party A from a constitutional contest, whereas the RHS of condition (1) is the expected contribution to the utility of Party A from a civil conflict. Condition (2) applies analogously to Party B.

Both condition (1) and condition (2) can be satisfied if and only if there exist quadruples, S , such that the sum, $P_A X_A + P_B X_B$, satisfies

$$(3) \quad Q_A + C_B/\beta > P_A X_A + P_B X_B > Q_A - C_A/\alpha.$$

If and only if either C_A or C_B are positive, then quadruples, S , that satisfy condition (3) exist. Hence, condition (3) has the following implications:

- Under the assumption of constitutional commitments, an incremental cost of civil conflict for either party would be sufficient to make it possible to design a constitution that is attractive to both parties.
- In a constitution that is attractive to both Parties, the sum, $P_A X_A + P_B X_B$ which equals the expected value of X , given that X is determined constitutionally, is neither

too large nor too small. In other words, the expectation of the constitutional outcome is neither too favorable or too unfavorable to either one Party or the other. The bounds on the expectation of the constitutional outcome are larger the smaller is either α relative to C_A or β relative to C_B .

Abide or abrogate?

In reality Parties cannot commit themselves to abide by the constitution.⁵ Moreover, without constitutional commitments the existence of quadruples, S , that satisfy condition (3) is not sufficient for constitutional resolution of a dispute to provide a viable alternative to civil conflict. As explained above, to be viable a constitution must be self enforcing.

Suppose that a polity establishes a constitution, described by a quadruple, S , that satisfies conditions (1) and (2). Suppose also that an initial constitutional contest takes place, and that Party A is the winner. Following this constitutional contest, but before Party A exercises its constitutional prerogative to choose X , if Party B finds the prospect of Party A exercising its constitutional prerogative to be unpalatable, then Party B can abrogate the constitution, leaving the dispute over X to be resolved by civil conflict. Assume that Party B will abrogate the constitution if and only if the present value of the contribution to its expected utility from abiding by the constitution would be smaller than the present value of the contribution to its expected utility from abrogating the constitution. To determine the present value of the contribution to its expected utility from abrogating the constitution, assume that, if the constitution is abrogated, then a civil conflict ensues in the current period and in every future period as long as the dispute about X recurs.⁶

⁵This assertion abstracts from the possibility that a Party could bond itself to abide by the constitution by offering collateral or other hostages. A large literature deals with the question of whether or not political actors can bond themselves. See, for example, Jeremy Bulow and Kenneth Rogoff (1989).

⁶An isomorphic assumption would be that the winner of a civil conflict in the current period, having subjugated the loser, would be able to choose X for the current period and for all future periods, except that under this assumption the incremental cost of a civil conflict would be incurred only once.

Given these assumptions, Party B, having lost the current constitutional contest, abides by the constitution, rather than abrogates the constitution, if and only if two conditions are satisfied. First, X_A is large enough to be credible, where credibility means that Party B expects that, if it does not abrogate the constitution, then Party A will abide by the provision or supplementary understanding that it will not set X larger than X_A . We analyze presently how large X_A must be.

Second, X_A is small enough that the quadruple, S , satisfies the following condition:

$$(4) \quad \beta (1 - X_A) + \frac{\rho}{1 - \rho} \beta [P_B(1 - X_B) + P_A(1 - X_A)] \geq \frac{1}{1 - \rho} (\beta Q_B - C_B).$$

The parameter ρ in condition (4) is an exogenous positive discount factor that, in addition to reflecting pure time preference, can depend on the hazard rate for termination of the dispute about X . The LHS of condition (4) is the present value of the expected contribution to the utility of Party B, the loser of the current constitutional contest, from abiding by the constitution in the current period and in future periods, given that X_A is credible and that Party B believes that Party A will not abrogate the constitution if Party B wins a future constitutional contest. The RHS of condition (4) is the present value of the expected contribution to the utility of Party B from abrogating the constitution, a choice that results in civil conflict in current and future periods. The specification of condition (4) assumes that Party B expects the parameters, β , ρ , Q_B , and C_B , to remain unchanged.⁷

If Party B does not abrogate the constitution, then Party A can exercise its constitutional prerogative to choose X for the current period. In choosing X , either Party A can abide by the provision or supplementary understanding that it will not set X larger than X_A or Party A can behave opportunistically, disregarding this limit and setting X equal to one, its most preferred value. Assume that Party A will abide by the limit on its constitutional

⁷We might suppose that, because Party A prefers X to be larger, if condition (4) is satisfied, then it is satisfied as an equality. The conclusions derived below do not depend on whether condition (4) is satisfied as an inequality or as an equality.

prerogative if and only if the present value of the expected contribution to its utility from abiding by the limit on its constitutional prerogative is at least as large as the present value of the expected contribution to its utility from behaving opportunistically.

If X_A is smaller than one, then Party A would get more current utility from opportunistic behavior than from abiding by the limit on its constitutional prerogative. Accordingly, Party A does not set X larger than X_A only if opportunistic behavior would incur a future penalty. To model this future penalty in a simple way, assume that, if Party A were to behave opportunistically, then Party B would infer that Party A is compulsively opportunistic. In other words, Party B would infer that Party A cannot make itself act in accord with the positive discount factor, ρ , but only can act instead as if the discount factor were zero.

Given this inference, if Party A were to show itself to be opportunistic, then for Party B the expected contribution to its utility from using a constitutional contest to resolve future recurrences of the dispute about X would not be larger than the expected contribution to its utility from a civil conflict. Thus, although it would be too late for Party B to do anything about the choice of X for the current period, Party B would not participate in constitutional contests in the future. Opportunistic behavior by the Party that won the constitutional contest would result in future recurrences of the dispute about X being resolved by civil conflict.

With Party A taking this consequence of opportunistic behavior into account, Party B takes X_A to be credible if and only if the quadruple, S , satisfies the following condition:

$$(5) \quad \alpha X_A + \frac{\rho}{1-\rho} \alpha (P_A X_A + P_B X_B) \geq \alpha + \frac{\rho}{1-\rho} (\alpha Q_A - C_A).$$

The LHS of condition (5) is the present value of the expected contribution to the utility of Party A from abiding by the limit on its constitutional prerogative in current and future periods. The RHS of condition (5) is the present value of the expected contribution to the utility of Party A from being opportunistic in current period and, thereby, provoking civil conflict in future periods. The specification of condition (5) assumes that Party A expects

the parameters, α , ρ , Q_A , and C_A , to remain unchanged. This formulation also uses the observation that the best opportunistic choice by Party A is to set X equal to one.

Can constitutional resolution provide a viable alternative to civil conflict?

Constitutional resolution of a dispute provides a viable alternative to civil conflict, regardless of which Party wins a constitutional contest, if and only if the quadruple, S , satisfies both condition (4) and condition (5) as well as the analogous conditions that apply if Agent B has won the current constitutional contest.⁸ Both condition (4) and condition (5) are satisfied if and only if the quadruple, S , satisfies

$$(6) \quad \frac{1}{1-\rho} (Q_A + C_B/\beta) \geq X_A + \frac{\rho}{1-\rho} (P_A X_A + P_B X_B) \geq 1 + \frac{\rho}{1-\rho} (Q_A - C_A/\alpha).$$

The analogous viability condition that applies if Agent B has won the current constitutional contest is

$$(7) \quad \frac{\rho}{1-\rho} (Q_A + C_B/\beta) \geq X_B + \frac{\rho}{1-\rho} (P_A X_A + P_B X_B) \geq \frac{1}{1-\rho} (Q_A - C_A/\alpha).$$

In condition (6) the sum, $X_A + \frac{\rho}{1-\rho} (P_A X_A + P_B X_B)$, represents the discounted present value of current and future realizations of X , conditional on Party A having won the current constitutional contest, assuming that both Parties will always abide by the constitution. In condition (7) the sum, $X_B + \frac{\rho}{1-\rho} (P_A X_A + P_B X_B)$, represents the discounted present value of current and future realizations of X , conditional on Party B having won the current constitutional contest, assuming that both Parties will always abide by the constitution. Thus, according to conditions (6) and (7), if constitutional resolution of a dispute provides

⁸With some effort we could derive the conditions under which constitutional resolution of a dispute provides a viable alternative to civil conflict for as long as Party A wins constitutional contests, but under which Party A would abrogate the constitution if Party B were to win a constitutional contest. Among possible examples of such a situation might be the Chilean constitution, which apparently was viable as long as the Right held political power, but which in 1973 the Right abrogated, when the Left, having gained political power, appeared to be behaving opportunistically.

a viable alternative to civil conflict, then the quadruple, S , is such that the discounted present value of current and future realizations of X , conditional on either one Party or the other winning the current constitutional contest, is neither too large nor too small.⁹ This property means that under a constitution that provides a viable alternative to civil conflict which Party wins the current constitutional contest does not matter too much to either one Party or the other.¹⁰

Conditions (6) and (7) do not uniquely determine the quadruple, S . Conditions (6) and (7), however, provide an answer to the prior question that this paper addresses: Is it possible to design a constitution that is self enforcing and that provides a viable alternative to civil conflict regardless of which Party wins a constitutional contest?

This answer to this question is yes if and only if there exists a quadruple, S , that would satisfy both condition (6) and condition (7). Furthermore, there exists a quadruple, S , that would satisfy condition (6) if and only if the parameters are such that

$$(8) \quad \rho \frac{C_A}{\alpha} + \frac{C_B}{\beta} \geq (1 - \rho) Q_B.$$

Also, there exists a quadruple, S , that would satisfy condition (7) if and only if the parameters are such that

$$(9) \quad \frac{C_A}{\alpha} + \rho \frac{C_B}{\beta} \geq (1 - \rho) Q_A.$$

Thus, it is possible to design a self-enforcing constitution only if the parameters satisfy both condition (8) and condition (9).

⁹If condition (4) and the analogous condition that applies if Agent B has won the current constitutional contest are satisfied as equalities, then the first weak inequality in condition (6) and the second weak inequality in condition (7) are equalities.

¹⁰If the second weak inequality in condition (6) is satisfied, then condition (1) also is satisfied. In addition, if the first weak inequality in condition (7) is satisfied, then condition (2) also is satisfied. Thus, if it is possible to design a self-enforcing constitution, then it would be possible to design a constitution that with constitutional commitments would be attractive to both Parties.

Conditions (8) and (9) imply that the following structural factors help to make constitutional resolution of a dispute a viable alternative to civil conflict:

- The parameters, α and β , which calibrate the importance of the dispute about X , are small relative, respectively, to C_A and C_B , the expected incremental costs of civil conflict.
- Neither Q_A nor Q_B is too large. This property means that neither Party has a big advantage in civil conflict.
- The discount factor, ρ , which measures the importance that the Parties attach to future outcomes, is large.¹¹

To illustrate the relevance of this theoretical analysis we use these results to explicate two dramatic historical examples of constitutional failure.

Why Secession? The Received Answer

In the course of American history the most consequential disputes between constituent groups of the polity involved slavery, the issue that divided Northern and Southern interests in ante-bellum United States.¹² The disputes over slavery are especially interesting because for seventy years artful constitutional compromises enabled Northern interests and

¹¹Conditions (8) and (9) imply that, if C_B/β is as large as Q_B and C_A/α is as large as Q_A , then a positive discount factor is not necessary for the existence of a self-enforcing constitution. But, conditions (6) and (7) imply that, if, as ρ approaches zero, the set of self-enforcing constitutions remains nonempty, then in the remaining self-enforcing constitutions X_A approaches one and X_B approaches zero.

¹²Postulating a polarity between the North and the South over the issue of slavery is a crude, but useful, simplification. In fact, neither Northern nor Southern interests were monolithic. Nevertheless, according to James McPherson (2001), “Since the 1950s most professional historians have come to agree with Lincoln’s assertion that slavery ‘was, somehow, the cause of the war’.” McPherson convincingly debunks the claim that the main Southern interest was not in defending slavery, but in “a noble cause, the cause of state rights, constitutional liberty, and consent of the governed.”

Southern interests to resolve these disputes by constitutional means. Indeed, for most of the ante-bellum period the understood prerogatives of winners of electoral contests under the Constitution shielded national politics from the issue of slavery.¹³

Nevertheless, as Barry Weingast (1998, pages 167-168) points out, “Because the country was growing, each new generation had to renew the arrangements that began when the founding fathers created a system with strong constitutional protection for slavery.” In 1861, however, both the existing constitution and new attempts to renegotiate the constitution dramatically failed, and events culminated in the secession of eleven Southern states from the Union and the ensuing Civil War.

Why did the issue of slavery eventually result in civil conflict? I take the received answer, my account of which is largely based on Robert Fogel (1989), McPherson (1988, 2001), David Potter (1976), and Weingast (1998), to involve three main elements:

First, by the middle of the nineteenth century, as Potter (1976, page 93) explains, “The longstanding sectional equilibrium within the Union was disappearing and the South was declining into a minority status, outnumbered in population, long since outnumbered and outvoted in the House, and protected only by balance in the Senate.” But, neither the Compromise of 1850, which admitted California to the Union as a free state, while allowing settlers in New Mexico and Utah to decide, under the principle of “squatter sovereignty”, whether these territories should become free or slave states, nor the Kansas-Nebraska Act of 1854, which organized the Kansas and Nebraska Territories under the principle of squatter sovereignty, resulted in the admission of additional slave states, as maintaining balance in the Senate would have required. In addition, as Potter (1976, page 93) stresses, “There was not one slave territory waiting to be converted into another slave state, while all of the upper part of the Louisiana Purchase, all of the Oregon territory, and now all of the

¹³Prior to the establishment in 1854 of the Republican Party, the main political parties, Whigs and Democrats, had national constituencies, and the sectionally divisive issue of slavery was not central in the competition between the parties.

Mexican Cession stood ready to spawn free states in profusion.” With their failure to gain admittance of Kansas as a slave state it was clear that Southern interests had permanently lost the protection of balance in the Senate.

Second, prior to the election of 1860 every President has been either a Southerner or a Northerner who had significant Southern support. But, by 1860 more rapid population growth in the North than in the South allowed Abraham Lincoln, the candidate of the newly formed Republican Party, to be elected without carrying a single Southern state. This unprecedented development confirmed that under the existing rules Southern interests now had lost the protection of the Presidential veto.

Third, the free-soil platform of the Republican Party, which called for the prohibition of slavery in the territories, implied a new understanding about the prerogatives of winners of electoral contests under the Constitution and, in effect, rescinded the understanding that the Constitution excluded slavery policy from national politics.¹⁴ Although the Republican platform did not call for emancipation, the new president, Lincoln, as quoted by Potter (1976, page 427) and McPherson (1988, page 179), had denounced slavery as “morally wrong”, had stated that “this government cannot endure, permanently half slave and half free”, and had expressed his hope for the “ultimate extinction” of slavery. According to Fogel (1989, page 381), the Republicans were “determined to restrict slavery’s political and economic domination to guarantee that the federal government promoted northern interests and principles.” On the Southern side, according to McPherson (2001), “Jefferson Davis...justified secession as an act of self-defense against the incoming Lincoln administration, whose policy of excluding slavery from the territories would make ‘property in slaves so insecure as to be comparatively worthless,...thereby annihilating in effect property worth thousands of millions of dollars’.”

¹⁴Weingast (1998) argues that this understanding depended on balance in the Senate and, hence, that the rescinding of this understanding was not an independent development, but rather a result of the increasing dominance of Northern interests in constitutional contests.

It is interesting to observe that, according to this account, in abrogating the Constitution the Southern secessionists were reacting both to demographic developments and to the proactive stance of the Republican Party in rescinding an implicitly understood limitation on the prerogatives of the winner of a constitutional contest.

We can easily adapt this analysis to the language of our model. Let Party A denote Northern interests, the winner of both the battle over statehood for Kansas and the presidential election of 1860, and let Party B denote Southern interests. (Recall that the term “Party” as used here refers to a group with a common purpose and not necessarily to a political party.) Also, let X equal to one represent the free-soil policy that Northern interests favored, and that Southerners saw as leading to destruction of the wealth of slave owners, and let X equal to zero represent a policy of unrestricted property rights for slave owners, without geographical limitations, that Southern interests favored. Intermediate values of X could represent a more moderate set of policies, which might include some geographical limitations on the property rights of slave owners and/or emancipation with compensation to the slave owners.

Using the language of our model, the first two elements in the received answer imply an increase in P_A , and the third element implies an increase in X_A . Apparently, these increases in P_A and X_A were so large that condition (4) was no longer satisfied. Accordingly, the existing constitution was no longer self enforcing, as Party B (Southern interests) found that the present value of its expected utility from abiding by the Constitution would be smaller than the present value of its expected utility from abrogating the Constitution.

Why Secession? A Deeper Analysis

The problem with this received answer is that it does not go far enough. Specifically, although the received answer implies a failure of condition (4) and, hence, can account for the failure of the existing constitution, it does not imply that condition (8) was no longer satisfied. But, if condition (8) was still satisfied, then the set of self-confirming constitutions was not empty. In that case it would have been possible to avoid civil conflict by revising

the constitution either explicitly or implicitly. The required revisions either would have changed the form of constitutional contests to reverse the increasing political dominance of Northern interests or would have constructed a new understanding limiting the prerogatives of Northern interests, as the likely winner of future constitutional contests. Moreover, given their experience in devising compromises, Northern and Southern interests should have been capable of realizing such a renegotiation of the constitution, if it were feasible. Certainly, there were many ideas for another compromise in the air.

One idea was to reconstitute the Union as a federation of the set of Northern states and the set of Southern states. In his proposal for a “concurrent majority”, Calhoun envisaged a dual presidency, with one president representing the North and one representing in the South, and each with the power to veto legislation. Of course, such a reform proposal had no chance, as Northern interests, having worked hard to destroy sectional balance in the Senate, would hardly be willing to accept a sectionally balanced presidency.

Another possibility would have been to limit the prerogatives of Northern interests to a policy choice no more extreme than the British example of emancipation with compensation. But, Fogel (1989, page 412) tells us that “whatever the opportunity for a peaceful abolition of slavery before 1845, it surely was nonexistent after that date. To Southern slaveholders, West Indian emancipation was a complete failure...They could see plainly that the economy of the West Indies was in shambles, that the personal fortunes of the West Indian planters had collapsed, and that assurances made to these planters in 1833 to obtain their acquiescence to compensated emancipation were violated as soon as the planters were reduced to political impotency.” In terms of our model, this account implies that, even if a limitation on X_A , the prerogatives of Northern interests, to a policy choice no more extreme than compensated emancipation would have satisfied condition (4), Southern interests did not view such a limitation to be credible.

The proposed Crittenden Compromise, perhaps the most serious of several futile attempts to amend the Constitution in order to prevent civil conflict, embodied another set of

possibilities for limiting the prerogatives of Northern interests. The Crittenden Compromise, formally introduced in Congress in December 1860, would have given explicit constitutional protection to slavery in those states, and in the District of Columbia, where slavery already existed and in those remaining territories in which slavery was to be allowed according to the Missouri Compromise of 1820. Both Northern and Southern interests rejected this compromise. The Republicans, led by President-elect Lincoln, would not accept any scheme that infringed on the free-soil plank of their platform. And, according to Fogel (1989, page 413), the Southerners by then “were convinced that northern hostility to slavery precluded a union that would promoted [Southern] economic, political, and international objectives.”

Finally, Northern interests might have accepted the establishment of an independent Southern Confederacy. Assuming that the Confederacy would have no territorial ambitions beyond the borders of the eleven secessionist states, such a peaceful dissolution of the Union would have allowed Northern interests to implement their free-soil policy in the territories. But, the fervent opposition of Southern interests to the exclusion of slavery from the territories belies this assumption. As Roger Ransom (1989, page 167) emphasizes, “The South of the mid-nineteenth century was an expansionist system that coveted land to the west and to the south...If they gained status as an independent nation, slaveowners would be free to pursue a ‘foreign policy’ just as inimical to the North’s interests as that pursued by the ‘slave power’ when it had control of the federal government within the union.” And, an independent Confederacy, unconstrained by the interstate commerce clause of the Constitution, would have had enhanced strategic advantages, including, for example, the ability to control access to the sea via the Mississippi River. Fogel (1989, page 416) argues that acceptance of an independent Southern Confederacy would only have postponed a war over slavery and its expansion and “that the delay would have created circumstances far more favorable to a southern victory.”¹⁵

¹⁵Massimo Bordignon and Sandro Brusco (2001) analyze the optimality of including secession rules in the constitution of a federal union. They consider a potential dispute over the value of the federal union. In

Why was civil conflict unavoidable?

This analysis suggests that in 1861 Northern and Southern interests failed to avoid civil conflict not only because the existing constitution was no longer self enforcing, but also, and more importantly, because constitutional resolution of the dispute over slavery no longer provided a viable alternative to civil conflict. In other words, the fatal fact was not that the existing constitution no longer satisfied condition (4), but that the perceived parameter values no longer satisfied condition (8).

What made civil conflict, which was avoided before the election of 1860, unavoidable in 1861? Historical scholarship, when combined with our theoretical model, suggests that the critical change was that, as a result of developments in the years leading up to 1861, the dispute between Northern and Southern interests became too important to be resolved through a constitutional political process. In terms of our model historical scholarship suggests that an increase in either α or β or both upset condition (8).

Fogel's account of northern ante-bellum politics suggests a plausible story that is consistent with an increase in α in the years leading up to 1861. From the late 1840s, mainly because of increased immigration, incomes and living conditions of native, northern, non-farm workers became increasingly depressed. Fogel (1989, page 356) tells us that this depression of living conditions was "one of the most severe and protracted economic and social catastrophes of American history."

As a consequence of this working-class depression land policy became increasingly important. Free homesteads, opening western lands for settlement by the working poor, became a paramount demand of northern labor. But, the objective of Southern interests that new territories be opened to slavery stood in the way of free homesteads. Thus, as Fogel (1989, page 350) explains, land policy "drew into direct conflict with Slave Power the northern working-class leaders who had previously remained aloof from the anti-slavery movement."

their analysis, in contrast to present analysis of the dispute over the expansion of slavery, secession resolves the dispute.

The result was the coalescing of free soil and nativist factions into the new Republican Party.

In addition, the evidence about the economics of slavery, as summarized by Fogel (1989) and Ransom (1989), suggests a plausible reason for an increase in β in the years leading up to 1861. According to Fogel (1989, page 412), “From the mid-1840s on...the slave economy of the South was vigorous and growing rapidly. Whatever the pessimism of [slave owners] during the economic crises of 1826-1831 and 1840-1845, during the last half of the 1840s and most of the 1850s they foresaw a continuation of their prosperity and, save for the political threat from the North, numerous opportunities for its expansion. The main thrust of cliometric research has demonstrated that this economic optimism was well founded...” As Ransom (1989, page 47) puts it, “On the eve of the Civil War, American slaveholders were coming off a decade and a half of exuberant growth and expansion.”

As it turned out, the actual incremental costs to both Northern interests and Southern interests of the ensuing civil conflict, including six hundred thousand men killed and thousands more maimed, certainly were larger than the expected incremental costs, C_A and C_B . We can speculate whether, if both Northern interests and Southern interests had not underestimated the incremental costs of the ensuing civil conflict, condition (8) still would have failed to be satisfied, even with the increased importance of the dispute, as reflected in increased values of α and β . But, it is only hindsight that suggests that a constitutional resolution, facilitated by a renegotiation of the constitution, would have been better for both Northern interests and Southern interests than the actual consequences of the civil conflict. In the event, the Civil War resolved the dispute. In the poignant words of Potter (1976, page 583), “Slavery was dead; secession was dead; and six hundred thousand men were dead.”

The National Socialist Revolution

In 1919 the Weimar Constitution was adopted to replace the abolished monarchy in the wake of Germany’s defeat in the Great War. The Constitution established a hybrid combination of parliamentary and presidential democracy. This hybrid provided two different political processes for resolving disputes among constituent groups of the polity. The usual

process was to be a parliamentary election and the creation of a parliamentary government, whose acts were subject to presidential veto. But, in an “emergency” the President could appoint a presidential government and authorize it to govern by decree, in effect suspending the parliament (Reichstag) and assuming dictatorial powers for himself.

As seems clear from historical accounts — see, for example, William Shirer (1960) — the Weimar constitution faced an impossible task in resolving disputes that reflected deep divisions among economic and social interests. In German society and politics workers, themselves split between Communists and Social Democrats, were aligned against employers, landless peasants were aligned against Junker landlords, monarchists were aligned against republicans, and nationalists and revanchists were aligned with the military against real or imagined subversives, whom, as Shirer (1960, page 157) tells us, the *Deutsche Zeitung* declaimed as “internationalist traitors and pacifist swine”.

These divisions resulted in political turmoil and other symptoms of social disintegration: widespread strikes, extremist political movements, and the formation of private paramilitary forces, which fought pitched battles in the streets. One consequence of the turmoil was indecisive economic policy, with a resulting hyperinflation, and, later, an inability to mount an effective response to the world-wide depression. These economic problems caused a further deepening of economic and social divisions.

The National Socialist German Workers’ (Nazi) Party, with Adolf Hitler as its leader, was founded in the midst of this chaos. The Nazi Party was a populist, nationalist, revanchist, and anti-Semitic movement that purported to bridge the divide between the left and the right of the political spectrum. The Nazis promised that, once they were in power, they would reestablish social order and take decisive action to solve economic problems. The Nazis emerged as an important political force for the first time in the parliamentary election of 1930.

Did the Nazis seek political power mainly to further a set of economic and social goals? Or, were they mainly opportunists who sought political power as a means to wealth and

personal aggrandizement? In either case their behavior reveals that the Nazi leaders were exceptionally ambitious men who put an unusually high value on achieving their objectives.¹⁶ Accordingly, using the language of our model, let Party B denote the Nazis, and let X equal to zero, the outcome most favored by the Nazis, represent a Nazi dictatorship that would be unconstrained in carrying out the Nazi program. Also, let Party A denote all groups that were opposed to the Nazis, and let X equal to one, the outcome most favored by those who opposed the Nazis, represent the exclusion of the Nazis from political power.

Four times between March 1932 and March 1933 the Nazis attempted, but failed, to win a constitutional contest for political power. In the first attempt Hitler ran for president. He campaigned vigorously, promising, as Shirer (1960, page 159) tells us, “jobs for the workers, higher prices for the farmers, more business for the businessmen, and a big Army for the militarists.” But, Hitler finished a distant second to President Hindenberg in both the initial vote on March 13th, 1932, and in the runoff election of April 10th.

The second attempt was in the parliamentary election of July 31st. The Nazis had maneuvered to bring about the dissolution of the Reichstag with the hope that a new election would sweep them into power. According to Shirer (1960, page 166), “the Nazis threw themselves into the campaign with more fanaticism and force than ever before.” Nevertheless, the results left the Nazis, although with 37 per cent of the vote the largest party in the Reichstag, far short of a even a simple majority.

The third attempt was in the parliamentary elections of November 6th. Again the Nazis had maneuvered to bring about the dissolution of the Reichstag. But, this time an increasingly widespread view that the Nazis were socially disruptive turned off both financial supporters and voters and, in turn, demoralized party comrades. The Nazis lost votes and seats in the Reichstag.

¹⁶Regarding the rank-and-file Nazis, according to Shirer (1960, page 206), “Most of them had belonged to the ragged army of the dispossessed and the unsatisfied. They...believed that the revolution that they had fought by brawling in the streets would bring them loot and good jobs.”

The final attempt was in the parliamentary election of March 5th, 1933. Social disintegration had led to what Shirer (1960, page 185) describes as “the inexplicable weakness, that now bordered on paralysis, of existing institutions — the Army, the churches, the trade unions, the political parties — [and] of the vast non-Nazi middle class and the highly organized proletariat.” In this setting, on January 30th, 1933, President Hindenberg had allowed the Nazis, in coalition with the Nationalists, to form a parliamentary government with Hitler as Chancellor. According to Shirer (1960, page 5), Goebbels, one of Hitler’s lieutenants, wrote that night in his diary, “The new Reich has been born. Fourteen years of work have been crowned with victory. The German revolution has begun.” But, Hitler’s new government had the support of only a minority in the Reichstag. Also, the Nazis held only three of eleven posts in the cabinet.

Yet again the Nazis, confident of overwhelming victory in a new election, maneuvered to bring about the dissolution of the Reichstag. Again, as Shirer (1960, page 189) tells us, “Goebbels was jubilant. ‘Now it will be easy,’ he wrote in his diary on February 3, ‘to carry on the fight, for we can call on all of the resources of the State. Radio and press are at our disposal. We shall stage a masterpiece of propaganda. And this time, naturally, there is no lack of money.’”

In addition to mounting a forceful election campaign, Hitler, on the pretext of a threat of a Communist revolution, prevailed on President Hindenberg to allow him to suspend civil liberties. Hitler also got Goering, another of his lieutenants, into the post of Prussian Interior Minister, whereby the Nazis gained control of the powerful Prussian state police. The police, supplemented by Nazi storm troopers, arrested political opponents and banned their publications. Despite these tactics, the election of March 5th gave the Nazis and their allies only a slim majority in the Reichstag, far less than the Nazis needed to amend the Constitution and to establish a dictatorship by constitutional processes.

In this election, although the Nazis used tactics that violated the spirit of democracy, they still had adhered to the letter of the Weimar Constitution. But, the results of the election

confirmed that the Nazis had little of no chance of achieving their goal of a Nazi dictatorship through electoral competition. Faced with the unpalatable prospect of their opponents exercising constitutional prerogatives to block the Nazi program, the Nazis, although still operating under a veneer of legality, abrogated the Weimar Constitution.

The Nazis won the ensuing civil conflict so quickly and easily that observers have blamed the opposition, perhaps unfairly, for giving up without a fight. On March 23rd, 1933, with a large number of opposition deputies detained by the police, in violation of the constitutional provision of legislative immunity from arrest, and with Nazi storm troopers lining the aisles, the Nazis rammed through the Reichstag an “Enabling Act” that gave dictatorial powers to Chancellor Hitler. Hitler then moved quickly to destroy any potential opposition. Within a few months, state legislatures were dissolved, and Nazi governors were appointed in each state; all political parties, other than the Nazi party, were banned; labor unions were dissolved; freedom of speech and press were abolished. In late June 1934 Hitler consolidated his personal power, and appeased the Army and his supporters on the Right, by carrying out a bloody purge of leftist elements among the Nazis. Finally, a month later, on the death of President Hindenberg, Hitler abolished the office of President, and extracted a personal oath of loyalty from the members of the armed forces.

In terms of our model why were the Nazis and their opponents unable to resolve their dispute over political power according to a constitutional political process? It is easy to see why the Weimar Constitution failed. The historical account suggests that Q_B , the probability of the Nazis winning a civil conflict, was large and that C_B , the incremental cost for the Nazis of a civil conflict was small. Given these parameters, P_B , the probability of the Nazis winning a political contest under the Constitution, apparently was too small to satisfy condition (4).

But, again we want to ask a deeper question. Why was condition (8), according to which constitutional resolution of a dispute can provide a viable alternative to civil conflict, not satisfied? As we have seen, if condition (8) had been satisfied, then it would have been

possible to avoid civil conflict either by revising the form of constitutional contests or by constructing, either explicitly or implicitly, a new understanding limiting the prerogatives of winners of constitutional contests. The historical account suggests that equation (8) failed to be satisfied in Germany in the early 1930s because, in addition to Q_B being large and C_B being small, the parameter, β , which calibrates the value that the Nazis attached to their objective of a monopoly of political power, was too large. Like the dispute over the expansion of slavery in ante-bellum United States, the disputed demand for a Nazi dictatorship was too important to be resolved by a constitutional political process.

Summary

Constitutional resolution of disputes among the constituent groups of the polity is desirable because it avoids the incremental costs of civil conflict. The theoretical analysis in this paper implies that it is possible to design a self-enforcing constitution that provides a viable alternative to civil conflict if and only if (1) no group regards the dispute to be resolved to be too important relative to the expected incremental cost of civil conflict and (2) no group has too big of an advantage in civil conflict. The theoretical analysis also implies that a constitution that provides a viable alternative to civil conflict limits the prerogatives of winners of constitutional contests such that the outcomes of constitutional contests for political power do not matter too much.

The paper illustrated the relevance of the theoretical analysis by applying these propositions to two dramatic historical examples of constitutional failure: the secession of eleven Southern states from the Union in 1861 and the National Socialist revolution in Germany in 1933. Historical accounts suggest that civil conflict was unavoidable in the American case because the dispute over the expansion of slavery had become too important, and perhaps also because the parties to the dispute underestimated the incremental costs of civil conflict. Historical accounts suggest that civil conflict was unavoidable in the German case because the disputed demand for a Nazi dictatorship was intrinsically too important and because, at least for the Nazis, the incremental cost of civil conflict was small.

Both of these examples illustrate an unfortunate reality. As long as the expected incremental costs of civil conflict are not too large, constitutional political processes cannot resolve important disputes. If the constituent groups of a polity are deeply divided and, hence, are unwilling to accept meaningful restrictions on the prerogatives of winners of constitutional contests, then civil conflict is unavoidable.

References

- Acemoglu, Daron and Robinson, James. "A Theory of Political Transitions." *American Economic Review*, 97(1), September 2001, 938-963.
- Alesina, Alberto. "Credibility and Policy Convergence in a Two-Party System with Rational Voters." *American Economic Review*, 78(4), September 1988, 796-805.
- Alexander, Gerard. *The Sources of Democratic Consolidation*. Ithaca, NY: Cornell University Press, 2001.
- Bordignon, Massimo and Brusco, Sandro. "Optimal Secession Rules." *European Economic Review*, 45(10), December 2001, 1811-1834.
- Bulow, Jeremy and Rogoff, Kenneth. "A Constant Recontracting Model of Sovereign Debt." *Journal of Political Economy*, 97(1), February 1989, 155-178.
- Dixit, Avinash, Grossman, Gene M., and Gul, Faruk. "The Dynamics of Political Compromise." *Journal of Political Economy*, 108(3), June 2000, 531-568.
- Fogel, Robert William. *Without Consent or Contract: The Rise and Fall of American Slavery*. New York: Norton, 1989.
- Gershenson, Dmitriy and Grossman, Herschel I. "Civil Conflict: Ended or Never Ending?" *Journal of Conflict Resolution*, 44(6), December 2000, 807-821.
- Grossman, Herschel I. "Kleptocracy and Revolutions." *Oxford Economic Papers*, 51, April 1999, 267-283.
- McPherson, James M. *Battle Cry of Freedom*. New York: Oxford, 1988.
- McPherson, James M. "Southern Comfort." *The New York Review of Books*, 48(6), April 12, 2001, 28-32.
- Potter, David M. *The Impending Crisis: 1848-1861*. Edited and completed by Don E. Fehrenbacher. New York: Harper and Row, 1976.
- Przeworski, Adam. *Democracy and the Market*. New York: Cambridge University Press, 1991.

- Przeworski, Adam. "Democracy as an Equilibrium." unpublished, March 2001.
- Ransom, Roger L. *Conflict and Compromise: The Political Economy of Slavery, Emancipation, and the American Civil War*. New York: Cambridge University Press, 1989.
- Shirer, William L. *The Rise and Fall of the Third Reich: A History of Nazi Germany*. New York: Simon and Schuster, 1960.
- Weingast, Barry R. "Political Stability and Civil War: Institutions, Commitments, and American Democracy", in Robert H. Bates, et. al., *Analytical Narratives*. Princeton: Princeton University Press, 1998.