THE AUTOCRAT'S CREDIBILITY PROBLEM AND FOUNDATIONS OF THE CONSTITUTIONAL STATE

by Roger B. Myerson

September 2006, revised August 2007

Abstract. A political leader's temptation to deny costly debts to past supporters is a central moral-hazard problem in politics. This paper develops a game-theoretic model to probe the consequences of this moral-hazard problem for leaders who compete to establish political regimes. In contests for power, absolute leaders who are not subject to third-party judgments can credibly recruit only limited support. A leader can do better by organizing supporters into a court which could cause his downfall. In global negotiation-proof equilibria, leaders cannot recruit any supporters without such constitutional checks. Egalitarian norms make recruiting costlier in oligarchies, which become weaker than monarchies. The ruler's power and limitations on entry of new leaders are derived from focal-point effects in games with multiple equilibria. The relationships of trust between leaders and their supporters are personal constitutions which underlie all other political constitutions.

Author's address: Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637. Phone: 773-834-9071. Fax: 773-702-8490.
Email: myerson@uchicago.edu. URL: http://home.uchicago.edu/~rmyerson/
The author is the Glen A. Lloyd Distinguished Service Professor of Economics at the University of Chicago.

Final published version: Roger B. Myerson, "The autocrat's credibility problem and foundations of the constitutional state." *American Political Science Review* 102(1):125-139 (2008): https://doi.org/10.1017/S0003055408080076

How art thou a king but by fair sequence and succession? Now, afore God... if you do wrongfully seize Hereford's rights... and deny his off'red homage, you pluck a thousand dangers on your head, you lose a thousand well-disposed hearts, and prick my tender patience to those thoughts which honor and allegiance cannot think. William Shakespeare, Richard II

This paper analyzes a simple model of autocratic politics to show how the elements of constitutional government can develop from basic problems of trust in the relationships between political leaders and their supporters. To win power and hold it, a leader must be able to make credible commitments to his supporters and agents. But credibility requires some threat of adverse consequences if commitments are not fulfilled. So any political leader, even an autocrat, must be judged by those who support him in power. Thus, we may find the primary institution of constitutional government in the autocrat's court, where courtiers implicitly judge the leader even as they serve him. The standards of behavior that active political supporters collectively expect of their leader, if he is to keep their trust, become a primary constraint on the leader's actions and may be viewed as an informal personal constitution for him. The establishment of formal constitutional structures in a state may depend critically on their compatibility with such personal constitutions for the leaders who hold high offices.

First and foremost, a successful leader needs a reputation for reliably rewarding his supporters. To compete for power under any political system, a leader needs the active voluntary support of many individuals, and these supporters must be motivated by some expectation of future reward in the event of their success. But when rivals have been defeated, a ruler may be able to enjoy the fruits of power without such broad support, and so an established ruler may be tempted to ignore the claims of past supporters. So the credibility of a leader's promises may be doubted if there is nothing that constrains rulers to fulfill their past promises (see Root, 1989, Shepsle, 1991, and North, 1993). Thus, leaders who are subject to constitutional constraints can have a competitive advantage over pure absolutists, whose sovereign freedom after winning power may reduce their ability to credibly recruit supporters beforehand.

The central thesis of this paper is that, as a minimal constitutional structure, a strong leader needs a court or council where his active supporters can collectively judge his treatment of them. The standards of behavior that they expect of their leader are a kind of law or personal constitution for him, defined and enforced within his own faction, which he must uphold or forfeit their trust. Such personal constitutions underlie all other constitutions of society.

Thus, although the main model of this paper begins with a conquest story about origins of the state, with power won by victory in battle, our analysis will nonetheless identify an essential role for a kind of collective agreement in the establishment of a viable state. Here, however, the essential social contract is not between a leader and the general population, but is between a leader and the active supporters who help him to defeat his rivals for power. A leader's ability to negotiate such a contract with his band of supporters underlies the main solution concept of negotiation-proof equilibrium that we will analyze here.

There are relatively few modern game-theoretic models of autocratic politics, and more are needed. Basu (2000) has considered simple models of dictatorship, and his distinction between dyadic and triadic power is fundamental to our formulation of absolutism, where the absolute leader is immune to any third-party pressure. The problem of supporters trusting a new leader, which is central in our analysis, also plays a crucial role in the selectorate model of Bueno de Mesquita et al (2003, pp 104-126). An understanding of institutions as a solutions for dynamic moral-hazard problems can be found in the models of Acemoglu and Robinson (2006), but their starting point is conflict among predefined economic classes, rather than simple rivalry among leaders. Svolik (2006) has found that authoritarian leaders' vulnerability to their inner circle of supporters may be the primary determinant of the duration of autocratic regimes. Egorov and Sonin (2005, 2006) have analyzed norms that regulate dictators' relationships with their ministers and with defeated rivals. Myerson (2007) has considered the leader's problem of credibly motivating governors whose positions entail opportunities for corruption and rebellion. Here our focus is on the problem of credibly motivating the captains who put the leader in power.

In the long philosophical literature on the foundations of the state, our approach here follows most closely Xenophon's Education of Cyrus. This classic study began with questions, which are still asked today, about what causes transitions between different kinds of political systems: democratic, oligarchic, and autocratic. Recognizing the essential role of political leaders as entrepreneurs who establish new political regimes, Xenophon saw the nature of the state as depending on the nature of leadership, which may be simplest to see in a one-leader autocratic regime. Thus, he suggested, an inquiry into the foundations of political institutions could well begin with a detailed case study of how an outstanding leader, Cyrus the Great, established an autocratic political institution, the Persian Empire. According to Xenophon, Cyrus established himself as a great political leader by cultivating a reputation for generously rewarding his captains after victory. So the essential point of his story is that a successful leader needs a reputation for reliably rewarding those who work to put him in power.

Such historical narratives and dialogues were the media of theoretical analysis for ancient social philosophers, but today we can seek new insights by modern game-theoretic analysis. In essence, however, our conclusions here may not be very different from Xenophon's.

A model of contests for power

Let us consider an island in which political power yields some income R, a flow of taxes and rents per unit time, which can be consumed or allocated in any way by the ruler of the island. To become the ruler of the island, a leader must first defeat the previous ruler in battle. Then, to stay in power, the ruler must defeat similar challengers who arrive as a Poisson process with an expected rate λ. (So in any time interval of length t, the expected number of challenges is λt. In any short time interval of length ε, the probability of a challenger arriving is $1-e^{-\lambda\varepsilon} \approx \lambda\varepsilon$.)

A leader needs active supporters to defeat any rival or challenger. We may think of these supporters as captains who bring military units into battle (or as precinct captains who deliver votes in a district). Let $p(n|m)$ denote a leader's probability of winning when he is supported by n captains and his rival is supported by m captains. Adopting a standard assumption for such contests (Skaperdas, 1990), let us suppose that there is some positive constant s such that

$$p(n|m) = n^s/(n^s+m^s).$$

We will see that the most interesting cases occur when s is between 1 and 2, representing situations where competitive forces have some moderately increasing returns to scale. Although our interpretation of force sizes may suggest restricting n and m to be integers, the mathematical analysis here will be simplified by allowing such force sizes to be any nonnegative real numbers.

Let c denote the cost for each captain to support the leader in battle against a rival or challenger. The leaders and captains are all risk neutral and discount future pay at some rate δ.

In the analysis of this model, we will use a few basic facts about the Poisson arrival of challengers which are summarized by the following lemma.

Lemma. Given any point of time when there is no immediate challenge confronting the ruler, let T denote the waiting time until the next challenger arrives. Then T is an exponential random variable with mean $1/\lambda$, the expected discount factor for payoffs delayed by T is

$$E(e^{-\delta T}) = \int_0^\infty e^{-\delta t} \lambda e^{-\lambda t} dt = \lambda/(\delta+\lambda),$$

and a stream of income y from now until the next challenger arrives is worth

$$E(\int_0^T e^{-\delta t} y\, dt) = E[y(1-e^{-\delta t})/\delta] = y/(\delta+\lambda).$$

Now consider a leader who is supported by n captains, each of whom is promised some income y as long as the leader retains power in this island. For simplicity, let us will consider stationary scenarios where all rivals are expected to have support from m captains. In such an environment, let $U(n,y|m)$ denote the expected discounted value of a captain's payoff, at any point in time when there is no challenger. By the lemma, $U(n,y|m)$ satisfies the recursive equation

$$U(n,y|m) = y/(\delta+\lambda) + [\lambda/(\delta+\lambda)]\, p(n|m)\, [U(n,y|m)-c],$$

and so

$$U(n,y|m) = (y-\lambda c)/[\delta+\lambda-\lambda p(n|m)].$$

When a challenger arrives, just before battle, the captain faces an immediate cost of c which will be followed rewards worth $U(n,y|m)$ only in the $p(n|m)$-probability event that his leader wins. Thus, the captain's expected payoff before going into battle is

$$-c + p(n|m)U(n,y|m) = [p(n|m)y - c(\lambda+\delta)]/[\lambda+\delta-\lambda p(n|m)].$$

We assume that the captain has the alternative to avoid battle and get payoff 0, and so the captain is only willing to fight if this expected payoff is nonnegative, which holds iff

$$y \geq Y(n|m) = c(\delta+\lambda)/p(n|m).$$

That is, $Y(n|m)$ is the smallest income stream that the leader can promise to his n captains, when all rivals are expected to have the support of m captains.

In this scenario, the leader will get personal use of the residual income $R-ny$ as long as he remains in power. So the expected discounted value of the leader's payoffs, at any point of time when he is ruler of the island and faces no immediate challenger, is the quantity $V(n,y|m)$

that satisfies the recursive equation

$$V(n,y|m) = (R-ny)/(\delta+\lambda) + [\lambda/(\delta+\lambda)]p(n|m)V(n,y|m),$$

and so

$$V(n,y|m) = (R-ny)/[\delta+\lambda-\lambda p(n|m)].$$

On the eve of battle against such a rival, the leader's expected discounted value of payoffs is

$$W(n,y|m) = p(n|m)\, V(n,y|m) = p(n|m)(R-ny)/[\delta+\lambda-\lambda p(n|m)].$$

Now the critical question is whether the leader's promises to pay his captains are credible. When we are considering the foundations of the state itself, we cannot assume any outside agency to enforce contracts between the leader and his captains. But as the leader's power is derived from his captains' support in battle, and the leader could be punished by his captains withdrawing their support in the future. So the constraints on what a leader can credibly promise to pay his supporters must depend on the structure of his relationships with his supporters. In our analysis, we will consider four different kinds of structures for the relationship between the leader and his active supporters. These structures effectively constitute four kinds of simple political systems, which we may call absolute monarchy, monarchy with a weak court, monarchy with a strong court, and oligarchy. We formally define each in turn.

Absolute leaders who are subject to no third-party judgments

An absolute monarch is one who is released from the constraints of law. But laws are effective only to the extent that people are expected to punish their violations. So an absolute leader is one who can violate norms of behavior without fear that other people will punish him in any way. Should we take this definition to imply that an absolute leader should fear no adverse consequences for denying promised payments to the captains who had supported him in battle? Such a leader, whose ability to rule and to get support against future challengers would not be affected in any way by his cheating previous supporters, would be unable to commit himself to pay any costly rewards for support in battle, and so he would be unable to credibly motivate any supporters. If people have rational expectations then his original bid for power should have gotten no support.

To avoid such a trivial conclusion, let us moderate the definition of absolutism to allow that, if an absolute leader cheated a supporter, then the cheated individual himself may respond by withdrawing support in the future, but that there is no organized structure to involve anyone else in punishing the absolute leader when he cheats a supporter. That is, we may say that a leader is absolute when his relationships with all supporters are purely bilateral, as if his agents have no communication with each other, only with their leader. So in absolutism, each of the captains who could potentially support the leader against future rivals cannot observe or respond to any change in the leader's relationship with other captains.

With this concept of absolutism, we can now characterize the sizes of forces which the absolute leader could be rationally expected to pay. As always, we simplify the analysis by assuming a stationary expectation that every other rival would bring a force of m captains against this leader. Against m, a force of n captains is <u>feasible for an absolute leader</u> iff there exists some wage rate y such that

(1) $y \geq Y(n|m)$ and $V(n,y|m) \geq V(k,y|m)$ $\forall k \in [0,n]$.

Here the first inequality is the supporters' participation constraint. It says that the captains' promised income y must be at least the minimal amount $Y(n|m) = c(\delta+\lambda)/p(n|m)$ that is required to motivate their support in battle. The second inequality is the absolute leader's moral-hazard constraint. It says that, after winning power, when there is no immediate challenge, the leader should not be able to increase his expected payoff by paying only k of his n captains, where $k \leq n$. The other $n-k$ who are cheated of their pay would respond by withdrawing support in future battles, so that the leader's expected discounted payoff would become $V(k,y|m) = (R-ky)/[\delta+\lambda-\lambda p(k|m)]$. Our assumption of absolutism implies that, after such an unanticipated deviation from n to k, the remaining k captains would still serve in battle for promises of $Y(n|m)$, as they would not realize that their probability of winning future battles had fallen to $p(k|m)$. To satisfy the second inequality, such a reduction of his probability of defeating future challenges must be enough to make the leader want to retain all his n captains at the cost y.

The preceding analysis assumed that the absolute leader could not replace any of his n captains after forfeiting their trust. As we have seen, if the absolute leader could costlessly replace any captain then he would have no incentive to pay any of them. But when n and y satisfy condition (1), absolutist feasibility can still apply with independent recruiting of new captains, provided that the absolute leader must make an advance payment of y/λ to gain the confidence of any new captain before he supports in battle. To verify this fact, notice that any force n which maximizes $W(n,y|m)-ny/\lambda$ for the leader just before a battle would also maximize $V(n,y|m)$ after the battle, because

$$V(n,y|m) = R/(\delta+\lambda) + \lambda[W(n,y|m)-ny/\lambda]/(\lambda+\delta).$$

A captain's service in battle can be rationally motivated only by expectation of pay after battle, but such advance pay y/λ can be a costly signal that the leader has not hired more captains than he will want to retain after battle. (Ray and Ghosh, 1996, have analyzed similar uses of relationship-building costs to allow rational trust when partners are replaceable and histories of cheating are not communicated.) Even with such recruiting costs, however, the absolute leader still could not credibly retain any force n that does not satisfy the feasibility condition (1).

It will be useful now to define the functions $v(n|m)$ and $w(n|m)$ by the formulas

$$v(n|m) = V(n,Y(n|m)|m) = [R-nc(\delta+\lambda)/p(n|m)]/[\delta+\lambda-\lambda p(n|m)],$$

$$w(n|m) = W(n,Y(n|m)|m) = [p(n|m)R-nc(\delta+\lambda)]/[\delta+\lambda-\lambda p(n|m)].$$

That is, $v(n|m)$ is the leader's expected value when there is no an immediate challenger, and $w(n|m)$ is the leader's expected value when there is a challenger to be fought, if the leader's n supporters are paid the minimal wage $Y(n|m)$ that motivates their support in battle against rivals of strength m.

For the leader, one possible benefit of increasing the size of his force n is that his captains would then be willing to fight for a smaller wage $Y(n|m) = c(\delta+\lambda)/p(n|m)$, because their probability of winning $p(n|m)$ is larger. But the leader can realize this benefit only if he can credibly convince his captains that he will actually retain such a larger force. Otherwise, even a captain who has not been cheated may worry that the leader has trimmed his payroll by cheating others. This credibility requirement is always problematic under absolutism, when captains cannot monitor each others' relationships with the leader. The following proposition asserts that an absolute leader could always benefit by committing himself to retain a larger number of captains than he can credibly maintain under absolutism. All proofs are in the Appendix.

<u>Proposition 1.</u> If the number of captains $n>0$ and the captains' wage rate y satisfy the feasibility condition (1) for an absolute leader against forces of size m, then there exist larger $k > n$ such that $v(k|m) > V(n,y|m)$ and $w(k|m) > W(n,y|m)$, so that the leader would be better off with k captains who are paid the required wage $Y(k|m)$.

Communication among supporters in the leader's court

By Proposition 1, an absolute leader would prefer to relinquish his absolutism and create an institution by which he can commit himself to a larger group of supporters. Such an institution can be created in our simple model merely by adding communication among the captains. Throughout history, princes and other political leaders have maintained courts in which their most important supporters are regularly gathered together. From our perspective, such a princely court can serve as a forum to guarantee that all the major supporters would learn about the leader's failure to appropriately reward any one of them. That is, the leader's court introduces an essential aspect of communication into the game. Even the unexpected absence of one courtier can be a tacit signal to the others that there is a problem in his relationship with the leader.

In our simple model, there is always the possibility of an equilibrium in which the captains do not expect the leader to pay them and so will not support him in any future battles, and so the leader has no reason to pay them anything. In this <u>distrustful equilibrium</u>, the leader could enjoy the consumption of the entire revenue R until he is defeated without support by the next challenger. This prospect of consuming all of R until the arrival of the next challenger would give the leader an expected discounted payoff of $R/(\delta+\lambda)$.

Now suppose that a complaint of any captain in the court would cause the leader and his captains to switch to the distrustful equilibrium. Such an event would make the leader worse off than paying n captains the income y iff $V(n,y|m) \geq R/(\delta+\lambda)$. Given n and m, this inequality can be satisfied with some motivating wage $y \geq Y(n|m)$ iff it can be satisfied with $y = Y(n|m)$. Thus, we may say that a force n is <u>feasible for a leader with a weak court</u> against m iff

(2) $$v(n|m) \geq R/(\delta+\lambda).$$

When this weak-court constraint is satisfied, the leader is willing to pay his n captains the wage $Y(n|m)$ that can motivate their continued support, with the understanding that his failure to pay this amount to any of them would switch them all over to the distrustful equilibrium.

The communication of grievances against the leader can be made incentive compatible in a noncooperative equilibrium of the game played by the members of court, even if the evidence of cheating is directly observable only by the leader and the cheated captain. When the leader and a captain both know that the leader has cheated the captain, it can be equilibrium behavior for them to switch bilaterally to a distrustful scenario, where the leader stops paying the captain, and the captain stops supporting the leader and complains in court (at least tacitly, by refusing to attend). On the other hand, as long as cheating has not occurred, each captain has a positive expected payoff $U(n,Y(n|m)|m) > 0$ on the equilibrium path, whereas his payoff would drop to 0 in the distrustful equilibrium. Thus, the captain would complain in court only if his expected payoff had been driven to 0 by an unexpected act of cheating by the leader.

This court is called <u>weak</u> here, because it cannot itself generate a new challenger to depose the leader, but it can only guarantee that the leader will get no support when the next challenger arrives. It may be strange, however, to assume that the probability of a new challenge soon would not be increased greatly when everyone at court knows that the ruler cannot get support against such challenges. It might be more reasonable to assume that a general loss of confidence at court would cause an immediate challenge to the current ruler, in which case we may call the court <u>strong</u>. With a strong court, the leader's expected payoff in the distrustful equilibrium would go to 0, as he would be immediately deposed. So we may say that a force n is <u>feasible for a leader with a strong court</u> against m iff

(3) $\qquad v(n|m) \geq 0.$

Proposition 2 offers general bounds on what can be accomplished with a weak court. These bounds depend on an assumption that the contest parameter s in $p(n|m) = n^s/(n^s+m^s)$ is not too small. A small parameter s less than 1 would represent a situation where forces have decreasing returns to scale, which seems unrealistic in most kinds of battlefields. (If s were less than 0.5 then an army that faces an enemy four times stronger would have more than 1/3 probability of winning.)

<u>Proposition 2.</u> Suppose that the force size n is feasible for a leader with a weak court against rivals of size m. Then the fraction of revenue that the weak-court leader pays to supporters is bounded by the inequality $nY(n|m)/R \leq p(n|m)\lambda/(\delta+\lambda)$, and the number of supporters is bounded by $n \leq R\lambda p(n|m)^2/[c(\delta+\lambda)^2]$. If $n > 0$ and $s > 0.5$ then the size of rivals' forces is bounded by $m \leq M_0$, where $M_0 = [R\lambda(2s-1)^{2-1/s}]/[4s^2c(\delta+\lambda)^2]$.

M_0 here may be called the <u>bound of suppression against weak courts</u>, because a weak court cannot provide any support when the rivals' forces m are larger than M_0. This bound of suppression on m and the bound on n in Proposition 2 both become small when the expected rate of challenges λ is small. That is, if the anticipated rate of challenges is small, then the weak court can justify at most a small force, and this only if rivals are small.

The constraints (1) on an absolute leader imply

$$v(n|m) \geq V(n,y|m) \geq V(0,y|m) = R/(\delta+\lambda).$$

(In fact, the conditions $V(n,y|m) \geq R/(\delta+\lambda)$ and $V'(n,y|m) \geq 0$ with $y \geq Y(n|m)$ are necessary and sufficient for absolutist feasibility (1).) So any force n that is feasible for an absolute leader is also feasible for a leader with a weak court. Thus, the bounds of Proposition 2 also apply to the forces of absolute leaders.

We may say that a force size m is globally feasible for leaders of any kind (absolute, or with weak courts, or with strong courts) iff m is feasible against m for such leaders. So m is globally feasible for leaders with weak courts iff a leader with a weak court can credibly retain the support of m captains when all rivals are also expected to have m supporters.

Our next proposition shows that, with some generality, the constraint of a weak court is costly for a leader. That is, in his contest for power, the leader's prospects could be improved by organizing his supporters into a strong court that could remove him from power.

Proposition 3. Suppose that $s \geq 2/3$. If a force n is feasible against m for a leader with a weak court and $0 < n \leq m$ then $w'(n|m) > 0$. (Here a prime denotes differentiation with respect to the first argument n.) So if m is globally feasible for leaders with weak courts then $\text{argmax}_{k\geq 0}\, w(k|m) > m$ (that is, on the eve of battle any leader would actually prefer to be committed to some force size k that is larger than m).

Negotiation-proof equilibria

The function w in our model represents the leader's objective at the formative stage of the new state, when the leader is organizing the supporters of his bid for power. When m denotes the anticipated strength of his rivals for power, the new leader's expected payoff would be maximized by negotiating a coalition of strength n where n maximizes $w(n|m)$. Indeed, within the terms of our model, if a leader could renegotiate the size of his coalition at any crisis point when he has to defeat a direct rival for power, then the leader would choose n to maximize this function $w(n|m)$. If this maximum against m is achieved by m itself, then each leader's optimized coalition would verify the others' expectations about their rivals. Thus, we may say that m is a negotiation-proof equilibrium iff

$$w(m|m) = \max_{n\geq 0} w(n|m).$$

Such an equilibrium m must be globally feasible for leaders with strong courts, because the condition $w(m|m) = p(m|m)v(m|m) \geq w(0|m) = 0$ implies $v(m|m) \geq 0$. But Propositions 1 and 3 tell us that, as long as the force-scale parameter s is not very small, such an equilibrium cannot be feasible for absolute leaders or for leaders with weak courts. Indeed, the following proposition shows that, unless s is very small, an absolutist or a weak-court leader would be unable to credibly recruit any supporters against such a negotiation-proof equilibrium, because it is above the bound of suppression against weak courts.

Proposition 4. When $s \leq 2$, the negotiation-proof equilibrium is

$$m_1 = Rs/[c(4\delta+2\lambda+s\lambda)].$$

In this equilibrium, the fraction of the revenue R that is paid to supporters is $m_1 Y(m_1|m_1)/R = 2s(\delta+\lambda)/(4\delta+2\lambda+s\lambda)$. When $s \geq 0.763$, this equilibrium m_1 is greater than the bound M_0 from Proposition 2, and so an absolutist or a leader with a weak court could not get any support against this equilibrium.

When s=2, the negotiation-proof equilibrium m_1 is equal to $R/[2c(\delta+\lambda)]$, which is the greatest force that is globally feasible with a strong court, and so the leader's expected payoff in equilibrium is 0.

When s>2, finding equilibria requires us to consider randomized (mixed) strategies. As Baye, Kovenock, and de Vries (1994) have shown, such contests with s>2 have equilibria in randomized strategies which yield expected payoff 0 for the contestants. The dynamic model here differs from their one-period contest model because the prize of victory in our dynamic model includes, not only the given value $R/(\delta+\lambda)$ of revenue from this contest until the next challenger arrives, but also the endogenously-determined continuation value of competing against the next challenger. When s>2, however, the zero-profit condition eliminates the continuation value, and so an equilibrium of a one-period contest model with prize $R/(\delta+\lambda)$

becomes equivalent to a negotiation-proof equilibrium of our dynamic model. This zero-payoff result implies that an absolutist leader or a leader with a weak court cannot raise any positive force against the randomized negotiation-proof equilibrium with $s>2$. (If some $n>0$ were feasible for a weak-court leader against the randomized equilibrium $\tilde{m}$, so that his expected payoff after winning would be at least $R/(\delta+\lambda)$, then the leader's optimal expected payoff before battle could not be less than $Ep(n|\tilde{m})R/(\delta+\lambda) > 0$.) So the suppression of weak courts and absolutists by negotiation-proof equilibria, which we found for the case of $0.763 \leq s \leq 2$, can be extended to $s>2$.

Thus, with broad generality, competition for power tends to create equilibria in which a viable leader needs a strong court that can remove him from power. Without such an institutionalized check on the leader, he could not credibly raise the support he needs to compete for power. The strong court provides the forum in which supporters can coordinate against the leader, if he violates his obligations to his supporters as understood by the court.

Oligarchic equilibria

The analysis in the preceding section was based on an assumption that a new leader would optimally negotiate, not only the size of his supporting coalition, but also the level of his obligations to these supporters. With n supporters against rivals of strength m, the leader's optimal level of obligation is to pay his supporters the lowest income $Y(n|m)$ that is required to motivate their support in the contest for power. Before a battle, the leader wants his supporters to have confidence that the threat of subsequently losing their confidence in his court would deter him from ever withholding this required income from them when he is in power.

But if the court can compel the leader to pay $Y(n|m)$, what is to prevent the courtiers from extracting even larger payments from the leader? The answer is that the courtiers are in a game with multiple equilibria, in which each courtier wants to support the leader as long as he believes that the leader still trusts him individually and that the other courtiers are planning to support the leader. The action of the court against the leader depends on a shift of equilibrium in a game with multiple equilibria, and thus depends on the culture of the court (as Schelling, 1960, has argued; see the discussion of focal coordination below). Our definition of negotiation-proofness above assumed that a new leader, by negotiation speeches to his band of prospective supporters before their first battle for power, could coordinate focal expectations of their joint behavior on the equilibrium that is most favorable to himself.

An alternative assumption, however, is that the collective culture could be tacitly or explicitly formed along lines that are more favorable to the rank-and-file supporters in the coalition. So let us now consider the extreme opposite assumption, that the strong court will insist that the rewards of power should be shared equally among all the active supporters of the regime. Under this assumption, the leader would effectively vanish from our story and be replaced by an egalitarian oligarchy of captains. When rivals are expected to have strength m, each of n oligarchs would, before the battle for power, have an expected payoff $\Omega(n|m)$ such that

$$\Omega(n|m) = -c + p(n|m)U(n,R/n|m)$$
$$= -c + p(n|m)(R/n-\lambda c)/[\delta+\lambda-\lambda p(n|m)] = w(n|m)/n.$$

The leading formateurs of such an oligarchic coalition (who may be assumed to feel sure of their own membership even if n were changed) would prefer to choose the size of their coalition n to maximize this expected payoff $\Omega(n|m)$. So we may say that m>0 is an <u>oligarchic equilibrium</u> iff

$$\Omega(m|m) = \max_{n\geq 0} \Omega(n|m).$$

In a limiting sense, we may also say that there is <u>oligarchic equilibrium at 0</u> if

$$\Omega'(n|m) < 0 \text{ for all n and m such that } n \geq m > 0,$$

because this condition implies that oligarchs would always prefer their group should be smaller than the anticipated size of their rivals.

We previously assumed that the autocratic leader could recruit new captains by promising them the least income Y that covers their cost of giving support. Now we are assuming that, in an oligarchy, egalitarian norms within the oligarchy would require that any new captains must be given equal status and equal expected rewards with other oligarchs. So the cost of recruiting new supporters in an oligarchy is greater than in a monarchy. Thus oligarchs should generally prefer to form a smaller coalition than would be optimal for a monarch.

In our main characterization of oligarchic equilibria, we find that an oligarchic equilibrium may be much smaller than the negotiation-proof equilibrium among monarchies that we found in Proposition 4. Indeed, when the expected rate of challenges λ is small enough, the only oligarchic equilibrium may be at 0.

Proposition 5. When $1 < s \leq 2$ and $\lambda \geq \delta(2-s)/(s-1)$, the oligarchic equilibrium is

$$m_2 = R[(s-1)\lambda-(2-s)\delta]/[c(\delta+\lambda)s\lambda].$$

If $s<2$ then this oligarchic equilibrium satisfies $m_2 < m_1$, where m_1 is the negotiation-proof equilibrium for monarchs, but the ratio m_2/m_1 is increasing in λ/δ and s. If s=2 then $m_2=m_1$. On the other hand, if $s \leq 1$ or $\lambda < \delta(2-s)/(s-1)$ then there is an oligarchic equilibrium at 0.

The coincidence of the oligarchic equilibrium and the negotiation-proof equilibrium of monarchies for s=2 extends also to the case of s>2, where the same zero-payoff randomized equilibrium applies to both. (In such equilibria where the oligarchs' expected payoffs are 0, random changes in membership of the oligarchy need not cause any disputes among oligarchs.)

An example

It may be helpful to consider a simple numerical example. To be specific, let us consider an example with revenue R=90, discount rate δ= 0.05, expected rate of challenges λ=0.2, cost of supporting in battle c=5, and force-scale parameter s=1.5.

[Insert Figure 1 about here]

With these parameters, Figure 1 shows how the force n that maximizes a new leader's expected payoff $w(n|m)$ depends on his regime-type (absolutist, weak court, or strong court) and on the anticipated size of rivals' forces m. The optimal size of an oligarchy to maximize its members' expected payoffs $w(n|m)/n$ is also shown.

When m is small, any kind of leader would prefer to raise a greater force n>m. But as m increases, credibility constraints tighten on absolutists and weak-court leaders. For this example, an absolute leader could not credibly raise any positive force when $m \geq 12.03$, and the bound of suppression against weak courts is $M_0 = 16.13$. A strong court can raise positive forces against any $m \leq 38.1$, but the leader's optimal n matches the opposing forces m in a negotiation-proof equilibrium at $m_1 = 30$. Oligarchs would prefer a smaller force than would be chosen by a single leader with a strong court, but the ratio of their best responses approaches 1 as m gets larger. Against the negotiation-proof equilibrium of monarchies at m_1=30, oligarchs would prefer to restrict their number to n=22. The oligarchic equilibrium, where oligarchs' optimal forces match what they expect of their rivals, is at m_2=18.

The vertical segments of the absolutist and weak-court best response curves (at 12.03 and 16.13) are actually discontinuities, not segments of indifference. That is, against m at these points of discontinuity, the feasible force at the top of the segment (n=12.97 for the absolutist at m=12.03, or n=25.6 for the weak-court leader against m=16.13) would be strictly better than 0 for the new challenger, but only 0 is feasible against higher values of m. So a concept of global negotiation-proofness subject to absolutist or weak-court feasibility would not exist here.

[Insert Figure 2 about here]

Figure 2 examines more closely a leader's choice of his force size n against the negotiation-proof equilibrium $m_1=30$. Increasing n would increase the number of supporters who must be paid, but would also increase the leader's probability of winning $p(n|30)$, which in turn would decrease the income $Y(n|30) = c(\delta+\lambda)/p(n|30)$ that each supporter must get. As we have required in equilibrium, the leader's expected payoff $w(n|30)$ on the eve of battle against a challenger is maximized by a force of n=30. But when there is no immediate challenge to be faced, the leader would prefer to increase his expected payoff $v(n|30)$ by reducing his paid force to n=23.3. However, $\max_{n>0} v(n|30) = 108.6 < v(0|30) = R/(\delta+\lambda) = 360$, and so a weak-court leader could not raise any positive force against $m_1=30$.

The required income for each captain-supporter in this negotiation-proof equilibrium is

$$y = Y(n|30) = c(\delta+\lambda)/p(n|30) = 2.5 \quad \text{when} \quad n=30 .$$

But an absolute leader could reduce n without changing the pay y for his remaining captains. So even before a battle, an absolute leader would prefer to increase his expected payoff $W(n,2.5|30)$ by reducing his force to n=20.6. This deviation would not be feasible if the supporters anticipated it, however, because their required rewards increase as the reduced force yields a reduced probability of winning.

When the supporters can monitor their number in the leader's court, the total cost of paying supporters $nY(n|30)$ is a U-shaped curve. As a fraction of total revenue, the required pay for supporters $nY(n|30)/R$ is never below 0.787 here, and it is $30\times2.5/90 = 0.833$ when n=30 in the negotiation-proof equilibrium. Even with a strong court, however, feasibility (3) against $m_1=30$ requires that a positive force n must be between 6.24 and 48.4, because $n<6.24$ or $n>48.4$ would make the required wage bill $nY(n|30)$ greater than the total available revenue R for this example. Having $n>48.4$ would be infeasible because there would be too many supporters to pay, even at the low wages $Y<1.86$. Having $n<6.24$ would be infeasible because the supporters' required wages would become too great at $Y>14.4$.

So even if the leader could credibly commit himself to share the entire revenue among his supporters, it could not be rational for anyone to support this leader in the battle for power unless there are enough other expected supporters to make his force greater than 6.24. Nobody wants to support a leader who is considered unlikely to get support from enough others. Thus, the captains' decisions about whether to support the leader is a coordination game, in which many-supporting and none-supporting can both be rational equilibria.

[Insert Figure 3 about here]

Figure 3 illustrates how the expected rate of challenges affects the different incentives for systems of oligarchies and monarchies. When the rate of challenges λ is high, the conversion of strong courts into oligarchies may have relatively little impact. But when λ is low and external challenges are infrequent, oligarchies tend to be much weaker than monarchies.

[Insert Figure 4 about here]

Figure 4 shows a discrete approximation of the randomized equilibrium that applies when $s=3$, keeping all other parameters as in our above example. In each battle, the rivals are expected to choose their forces $\tilde{n}$ and $\tilde{m}$ independently according to this probability distribution. The winner, say with $\tilde{n}=n$ supporters, will have to pay each of them an income $c(\delta+\lambda)/Ep(n|\tilde{m})$ until the next challenger arrives. In this discrete approximation, we find a probability 0.79 at $n=43$, probability 0.15 at $n=10$, probability 0.04 at $n=3$, and probabilities 0.01 at $n=1$ and $n=0$. The actual equilibrium seems difficult to characterize, but we know that it has expected value

$$E(\tilde{n}) = E(\tilde{m}) = R/[2c(\delta+\lambda)] = 36,$$

so that the expected costs of supporters in battle fully exhaust the expected discounted value $R/(\delta+\lambda)$ of the revenues that the winner will get until the next challenger arrives.

Numerical analysis of other values of s shows that, as s increases, the equilibrium probabilities become scattered more evenly over the interval from 0 to $R/[c(\delta+\lambda)]=72$, approaching a uniform distribution on this interval as $s\rightarrow\infty$.

Focal coordination among multiple equilibria as a foundation of political power

In the above analysis, we assumed that a leader needs captains' costly support only in occasional battles against rivals, and that such battles are separated by peaceful periods when an incumbent ruler can reap fruits of power without the active support of all these captains. A defense of this assumption must be formulated in terms of some more fundamental model of how established rulers can exercise power in society. So in this section, to justify this crucial assumption that an incumbent ruler can govern without support of the captains who put him in power, we need to take another view on the origins of government.

Observing that most states had been founded by conquest, rather than by some original contract among the governed population, Hume (1748) argued that the foundations of political power generally depend, not on any prior consent of the population, but merely on a common recognition by the population. That is, the establishment of a sovereign government may be effected by a generally shared perception or belief of the population.

From a modern game-theoretic perspective, Schelling's (1960) focal-point effect explains how such shared perceptions can be socially decisive. Schelling argued that, in games that have multiple Nash equilibria, any cultural or environmental factor that focuses people's attention on one equilibrium can make it rational for everyone to act according to this equilibrium, as any single individual would suffer from deviating unilaterally. So by the focal-point effect, the effectiveness of one equilibrium instead of another is determined by a commonly shared understanding or belief. As Hardin (1989) has observed, severe costs of anarchy can make the process of constituting a state into a game with multiple equilibria.

To be more specific, let us consider a simple story about how focal coordination can provide the foundations of political power, and can be the effective source of the ruler's revenue R in our model. (See Myerson, 2004, for a fuller development of this story.) Let us suppose that the island which we have been considering is actually inhabited by a large population of peasants, who go out each day to harvest the fruits of fields all over the island. But when a peasant arrives at a ripe field, he generally finds another peasant nearby, and each must decide for himself whether to claim the harvest here, or defer. A peasant who defers gets payoff 0. If one peasant claims while the other defers, then the claimant gets a positive payoff $r > 0$. But if two matched peasants both try to claim then they will suffer a conflict that costs each player $-\kappa<0$. So each matched pair plays the following rival-claimants game (also known as chicken):

	Player 2 claims	Player 2 defers
Player 1 claims	$-\kappa, -\kappa$	r, 0
Player 1 defers	0, r	0, 0

This simple two-player game has three possible equilibria. There are two pure-strategy equilibria where one of the players claims and the other player defers, yielding the payoffs (r,0) or (0,r). The game also has a symmetric randomized equilibrium, in which each player has an independent probability $r/(r+\kappa)$ of claiming, and the expected payoffs for both players are 0 in this symmetric equilibrium.

For each of these local matches, a recognized social leader can make a focal selection among the equilibria by designating one of the players to have claiming rights in the match. The leader's suggestion that the designated player should claim, and that the other should defer, can become a self-fulfilling prophecy because, when each player believes that the other will act as the leader suggests, then each player finds that his own expected payoff is maximized by following the leader's suggestion. The player who has been granted claiming rights by the recognized leader can confidently claim to get $r>0$, and the other player should prudently defer to get $0>-\kappa$. There is no need for any external sanction to compel players to obey the leader, because each is motivated to obey the leader by the perception that the other will also obey him. Thus, as Hume suggested, the leader's power of jurisdiction here depends only on the general popular opinion or belief that he has this power.

Such focal coordination power could be vested in any individual, provided only that each matched pair must share a common understanding about who has jurisdiction over their case. So the problem of agreeing about social leadership remains as a social coordination problem, but it is the coordination problem to solve all other coordination problems. Focal factors that bestow such coordination power on a leader may be called legitimacy (or charisma when they are intrinsic to the leader's personality). By the focal-point effect, the selection of legitimate leadership in any society can depend on its particular culture and history, such as a local tradition of identifying a particular family as royal.

Such a position of focal leadership is worth fighting for, because the power to allocate valuable rights can be profitable. If the payoffs from these rival-claimants' games are in some transferable units like money, then a player should be willing to pay anything up to the value of the prize r for the leader's authorization to claim it. Such transactions might be deterred if the act of taking a bribe could cause a leader to lose his socially recognized position of power. But a leader whose status is secure would be in a position to take rents up to the value r from each rival-claimants' game that he decides each day on the island. Thus, the assumption in our main model that an incumbent ruler can earn a flow of rents R can be derived from these profits of allocating claiming rights to the peasants of the island.

In the absence of any general natural law for determining legitimate rulers, there would seem to be little that a theorist could say in general about the allocation of such power. But we can say more if we admit that rivals may sometimes compete for power and if we make a few broad assumptions about the nature of such competition. We may naturally assume that, although a general recognition of ruling status may allow a leader to hold uncontested power of focal coordination in a society over extended intervals of time, there will be some points in time when a challenger can try to take power, and then the incumbent will have to defend his status. The outcome of such a contest for power will be to install one of the rivals as the generally recognized holder of focal coordination power, and so the contest's outcome may be naturally influenced by the number of people who actively support each rival in the contest.

Thus, as a simple and broadly interpretable model of politics, we may naturally assume that the recognized ruler of the island at any point in time will be the person who, with his supporters, has most recently won a battle for leadership of the island. Our model can admit different interpretations of what a battle is, from the Battle of Hastings to a presidential election, provided only that the people of the island must share a common understanding of what these battles are, when they occur, and how the winner is recognized. We can assume that such battles occur only intermittently; but, when a battle occurs, a greater force of active supporters would increase a contender's chances of winning.

Thus, force may serve as a factor for determining political leadership, but the recognized leader has no further need of such force when he allocates claiming rights among the general population in their coordination games. As we have seen, claiming rights that have been granted by the recognized ruler can be self-enforcing simply by creating expectations of asymmetric behavior by members of the local population, and so armies of supporters may not be necessary for the ruler to exercise power between the intermittent battles.

A leader's personal constitution

Such focal-coordination theories suggest that the foundations of political institutions are a matter of multiple equilibria, rationally indeterminate until the focal effects of local cultural traditions are taken into account. From this perspective, let us return to the main contest model of this paper. In this model, by making some parametric assumptions about the frequency and intensity of contests for power, we have been able to derive mathematical conclusions which say something deterministic about the characteristics of a negotiation-proof equilibria among rivals for power. The determinacy of this analysis is only about the size of the supporting forces, however, not about who actually gets to be a leader. Any particular individual's status as a leader must depend on social interactions that admit multiple equilibria.

This multiplicity of equilibria is essential in the analysis of our contest model, because the leader's motivation to reward past supporters depends on the expectation that the leader could otherwise lose their support in the future. The exchange of a captain's support for a leader's promised rewards involves a relationship of trust. Such relationships of trust and distrust correspond to different equilibria of the same dynamic game. As we have seen, an individual captain should not give costly support to a leader who is not expected to reward the support, and a leader has no incentive to give costly rewards to past supporters if withholding these rewards would not reduce his expectations of future support. So there can always be an equilibrium where individuals outside of any arbitrary group would not trust or support the leader, and where the leader would never reward such individuals.

Furthermore, we have seen that an individual should not give any costly support to a leader who is fighting for power when it is expected that nobody else is likely to give support, because the leader's probability of winning would be so small that even credible promises to share the entire state revenue R would not be worth the individual's cost c of giving support. So there can always be an equilibrium in which an aspiring leader gets no support and so has no chance of winning a contest for power.

Thus, an individual's status as a leader depends on an equilibrium relationship of trust with a group of supporters, who are confident of the leader's reliability and of their own collective strength. In this equilibrium, the leader's reliability is made credible by the threat that any deviation could make them all switch to an alternative equilibrium in which people do not have such confidence in the leader. A reliable reputation with many active supporters is the rare and fragile asset that defines a leader.

So the relationship of trust between a leader and his active supporters, like any reputational equilibrium, can exist only as one of many possible equilibria. Thus, the creation of such leader-supporter relationships can be influenced by arbitrary cultural traditions, according to the focal-point effect. (Even in the wilds of ancient Germany, according to Tacitus [1970, p. 112], fighting men would serve in the retinue of a mere boy if he inherited a position of leadership.) The legitimacy or charisma that a leader needs to gather a confident army of supporters may be bestowed on individuals by random focal events, which may occur only at random points in time. Thus, the arrival of new challengers could indeed be a Poisson process with rate λ, as we assumed in our model.

Our results (Proposition 4 in particular) show how competitive pressures can create an environment in which a leader cannot credibly recruit any force of supporters without a strong court. That is, to maintain reliability for a larger force, the leader's supporters need a forum where they can communicate grievances, and they need a shared sense of group identity so that they will all react to a breach of trust against any one supporter.

Our model can be extended to allow that supporters may retire from court, after their service has been appropriately rewarded, and new supporters may be recruited. But any new supporter needs to be accepted by this group at court, to be assured that his mistreatment by the leader would cause them all to distrust the leader. In turn, the courtiers need to monitor and regulate the leader's recruiting, because a leader who could always freely recruit new supporters would have no incentive to pay his old supporters for their past service.

We have not assumed any costs of participating in the leader's court. But if there were costs of attending court, then participation in the leader's court could itself be required for a good relationship with leader. For example, traditional feudal oaths required a vassal to give aid and counsel to his lord. Support in battle would be aid, and counsel could mean regular participation in the lord's court, where his reputation with all vassals is maintained (Finer, 1997, p. 870).

The rules that define what a leader must do to maintain his supporters' trust may be regarded as a personal constitution for the leader, even when the leader rules as a autocrat who is not constrained by any formal written constitution. In this personal constitution, the leader is constrained to share benefits of power with a privileged group of supporters. To them at least the leader must give a kind of justice, or else lose their collective confidence and support.

The bounds of the court's protection have a certain arbitrary element that depends on the courtiers' shared expectations. On the one hand, if there are people whose identity is considered alien to the members of court, then the leader's reputation among his courtiers might not be affected by his cheating such aliens, who thus could not hold any protected claims on the leader. On the other hand, the leader may have other obligations of a different nature that are considered germane by the court, so that the leader's failure to meet one of these obligations would stimulate the same reaction at court as his cheating a prominent courtier. That is, the leader's reputation at court may also be used to enforce any other constraints on the leader's behavior that are recognized by the courtiers. It is only necessary that a violation of these constraints would be observed by the members of court and would shift their expectations to a distrustful equilibrium in which nobody has the confidence to support the leader.

In particular, a political leader may fear to violate the terms of a formal constitution if such a violation would seem to his supporters like cheating one of them. Such a linkage may be particularly natural if the leader regularly proclaimed obedience to this constitutional system while developing his relationship with supporters, so that its violation would be a shocking change from the pattern of behavior that the supporters have come to trust. Thus, the effective power of a formal constitution to constrain political leaders may be based on leaders' fundamental need to maintain a fragile relationship of trust with a group of supporters.

Implications for more complex political systems

Agency incentive problems are fundamental in any political system. Constitutional rules are enforced by actions of political leaders and government officials, who must be motivated by an expectation of rewards and privileges as long as they fulfill their constitutional responsibilities (as in Becker and Stigler, 1974; see also Myerson, 2007). So the survival of any political system depends on its providing appropriate incentives for political agents to take actions that may be subject to moral-hazard temptations and imperfect observability.

In particular, at the birth of a new regime or in a crisis that threatens the regime's survival, supporters of the regime must exert efforts that can be rewarded only in the future. Then a political leader must serve as a banker, whose promises of future credit are trusted and valued as rewards for current service. But when the crisis has passed, the leader's need for such support is reduced, and he may be tempted to withhold costly rewards for past support. Thus, to better understand the foundations of the state, we have considered a simple model of political systems that is designed to highlight the central moral-hazard problem at the highest level of politics: the leader's temptation to deny past promises to the supporters who put him in power.

This problem of trust is fundamental to the nature of political leadership and has broad consequences for the nature of the state. We have argued that a leader may be defined as someone who has a reputational equilibrium of trust with a strong group of supporters. Such a reputational equilibrium can be effective only if it is jointly believed by the individuals involved, which is essentially a question of multiple equilibria. So by the focal-point effect, an individual's status as trusted leader or trusted supporter is can depend on arbitrary cultural traditions. But we have argued that, regardless of the cultural background, a strong leader needs a joint reputation among his supporters, which depends on their regularly communicating with each other about their relationships with the leader. A purely absolute leader, whose treatment of past supporters is not subject to any third-party judgment, would have very limited ability to make credible promises to supporters before attaining power. So we found that a new leader's political prospects can generally be improved by subjecting himself to oversight in a court where his mistreatment of any past supporters could cause his own downfall. Indeed, we found competitive equilibria where a leader could not credibly recruit any supporters at all without an expectation of such collective oversight.

Our model is only a simplified abstraction of real political systems, but it offers insights into the nature of political leadership that can be applied to better understand the complex political systems of real life. Our main result, that a successful leader's power must be based on his reputation in a forum where supporters can communicate grievances, implies an essential generalization of the traditional English doctrine that sovereignty is held, not by a king alone, but by a king in parliament. The term *parliament* itself comes from a word for communication.

Our results should call into question simplistic notions about the absolute nature of monarchs throughout history. Indeed, Finer (1997, pp. 1307-1335) has observed that supposedly absolute monarchs like Louis XIV were actually highly constrained by traditional concepts of aristocratic rights and courtly privileges, as our model would predict.

There are, of course, historical cases where rulers like Ivan the Terrible and Stalin have tyrannically abused their own courtiers. By sowing seeds of suspicion that prevent others from coordinating against him, an established tyrant can trap a society in a general reign of terror that afflicts even his own past supporters. In the framework of our model, such a tyranny would correspond to an absolutist regime, where the captains are denied any possibility of coordinating against their leader. In our analysis, the competitive weakness of such absolutism is derived from the basic assumption that any regime must eventually face some external challenges, and at such times the regime's survival will require active voluntary efforts by its captains (who can be trusted to serve loyally and take crucial initiatives in chaotic battles only if they have sufficiently positive stakes in the regime). In particular, such voluntary support is essential in the initial movement to establish a new regime, when a new leader does not yet have power to motivate his supporters by threats of punishment alone. So a leader who was expected to subsequently rule as an arbitrary tyrant would be unable to recruit supporters for his original rise to power. Thus, although our simple model cannot pretend to explain all historical events, our analysis offers a reason to view such tyrannies as exceptional cases that necessarily involve a failure of rational expectations by early supporters of the regime. The regularization of secure privileges for the party elite in the Soviet Union after Stalin's death is closer to our model's predictions and may be closer to the normal pattern of authoritarian regimes in history.

Our insights about leadership may be applied to leaders other than the senior ruler of a state. The political leader's n captains in our model may themselves be leaders of their own supporting staff, who must similarly trust that their efforts will be rewarded by their captain. At each level, the leader-supporter relationships of trust are matters of multiple equilibria, and so the networks of trust that exist in a society may depend on arbitrary facts of its culture and history. The trustworthy status of a leader at any level can be buttressed by an institutional court or council where his supporters communicate. So a feudal system can be a hierarchy of leader-supporter (or lord-vassal or patron-client) relationships, where each leader has own court of supporters to maintain his reputation for appropriately rewarding them (Finer, 1997, p. 870). But outside of pure feudalism, for any intermediate leader who is not at the sovereign summit of the hierarchy, there is also a possibility that the leader's superior can help adjudicate disputes about the proper rewards for supporters.

Once we recognize that the size of a leader's circle of trusted supporters may be defined by exogenous social expectations, we can see a possible danger of usurpation if a subordinate captain enjoys the trust of a force that is disproportionately large in comparison with the supreme leader's circle of trust. Thus, a weak king who is not widely trusted may need ministers and generals who are similarly weak, as Egorov and Sonin (2006) have noted.

Our analysis suggests that the potential effectiveness of a clandestine political organization like al-Qaeda may depend critically on its having a secure base where activists can gather for joint communication with their leaders. If a leader could not hold court with his active supporters and had no system for regularly paying and auditing them, then he could not build the personal trust that is needed to motivate their costly efforts.

Our view of a leader's personal court as the foundation of constitutionalism may find some confirmation in the historical evolution of these institutions. Many great institutions of civilization were initially developed to unite rival princely courts or to extend guarantees to a wider circle of regime supporters. The development of the English common law around 1160 began with Henry II's need to assure equal treatment to the former supporters of both sides of the recent civil war between Stephen and Matilda (Warren, 1973, pp. 332-3). Similarly, the great development of the Chinese civil-service system around 980 served to guarantee a fair allocation of patronage jobs among former courtiers of the Ten Kingdoms which had accepted integration into the Song Empire (Bol, 1998, p. 188). Rights and privileges of these institutions were later extended to more of the population, but from the start their enforcement relied on an understanding that a ruler who violated them could lose elite supporters' trust.

The extension of political rights to broader groups of the population has, from ancient Greece to modern Europe, often been driven by military competition which compelled states to earn the loyalty and trust of larger forces in times of war (de Jouvenel, 1948, p. 7). The general principle is that costly efforts to support and defend a political regime cannot be expected from people who have no political power to enforce any claims for future rewards. But as we saw in the analysis of oligarchic equilibria, members of an elite oligarchy have a greater incentive to restrict the size of their politically privileged class, even at the cost of weakening their state, compared to a monarchy. Thus philosophers of the Enlightenment could find that equality before the law seemed more compatible with unconstrained monarchies than with political systems that distributed power across aristocratic institutions (Finer, 1997, pp. 1434-8).

In democracy, of course, political leaders are supposed to extend their base of support to include voting masses. Campaigns and elections are the extended court in which leaders are constrained and judged by the voters. But democratic constitutions must be upheld by the actions of elected political leaders, and an inner circle of active supporters, small enough to monitor, is as essential for political leaders in democracy as in any other political system. Competition for leadership within a political party is rarely regulated by external laws, and a successful party leader needs active supporters within the party to whom he may be bound only by the kind of personal constitution that has been discussed here. Of course the larger constraints of a democratic constitution should limit a successful leader's ability to divert public resources to reward his active supporters. But the voters' evaluation of a political leader should take account of the circle of active supporters around him, as his relationship of trust with these supporters is a primary political commitment for the leader.

In particular, a democratic constitution would be imperiled if its most powerful office were held by a leader who could be confident that his active supporters would still trust him after he openly violated the constraints of the constitution. This point may help us to understand the fundamental differences between long-established democracies and new democracies. In a nation with a long continuous history of democracy, senior political leaders have developed relationships with their active supporters in a general context of obedience to democratic constitutional constraints, and so deviations from democratic norms would be naturally seen as a shock that could jeopardize these relationships. But in a nation where constitutional democracy has not previously existed, senior political leaders may have developed their bases of support with regular use of tactics that are against the norms of constitutional democracy, such as private violence or corrupt diversion of public funds; and then it is hard to expect that such leaders would suddenly begin to fear that continuing these regular practices might jeopardize their relationships with active supporters. So the chances of long-term success for a new democracy may depend on allowing more independent opportunities for different leaders to begin cultivating new democratic reputations, which can be facilitated by a federal division of powers or a multiparty parliamentary system. (This idea is developed further by Myerson, 2006.)

More generally, the effective terms of a new constitutional government can be constrained by the nature of pre-existing political relationships. The rules of a new regime are not written on a blank slate. The first officials under a new constitution need support to win their high offices, and so they cannot be expected to abandon their past supporters at the start of the new constitutional system. Provisions of the new constitution would be unenforceable if they asked these leaders to violate the terms of longstanding relationships with supporters. Thus, the fate of a new constitution may depend critically on the pre-existing personal constitutions that bind its first political leaders with their primary supporters.

Appendix: Proofs of formal propositions

Proposition 1. If $n>0$ and y satisfy the feasibility condition (1) for an absolute leader against m, then there exist $k > n$ such that $v(k|m) > V(n,y|m)$ and $w(k|m) > W(n,y|m)$.

Proof. The absolutist constraints (1) imply that $V'(n,y|m) \geq 0$, because otherwise the leader could gain from a small decrease of n, holding the wage y fixed.

If $y > Y(n|m)$, then the proposition could be trivially verified by decreasing the wage to $Y(n|m)$, which yields strict gains for the leader, and then choosing k close enough to n for continuity of all expressions to maintain the strict inequalities. So now, to complete the proof, we consider the case where $y = Y(n|m)$, so that $V(n,y|m) = v(n|m)$ and $W(n,y|m) = w(n|m)$.

Notice $Y'(n|m) < 0$, because $Y(n|m) = c(\delta+\lambda)/p(n|m)$ and the probability $p(n|m)$ is increasing in n. So the total derivative of $v(n|m) = V(n,Y(n|m)|m)$ with respect to n is

$$v'(n|m) = V'(n,y|m) - Y'(n|m)n/[\delta+\lambda-\lambda p(n|m)] > 0.$$

Thus, any sufficiently small increase of force size from n to some $k > n$ will increase the expected leader's expected payoff v when there is no immediate challenge. When there is a challenger to be fought, the leader's expected payoff w is also increased because

$$w'(n|m) = p(n|m)v'(n|m) + p'(n|m)v(n|m) > 0.$$

Proposition 2. Suppose that n is feasible for a leader with a weak court against m. Then $nY(n|m)/R \leq p(n|m)\lambda/(\delta+\lambda)$ and $n \leq R\lambda p(n|m)^2/[c(\delta+\lambda)^2]$. If $n > 0$ and $s>0.5$ then

$m \le M_0$, where $M_0 = [R\lambda(2s-1)^{2-1/s}]/[4s^2c(\delta+\lambda)^2]$.

Proof. Writing $p=p(n|m)$ and $Y=Y(n|m)$ for short, the weak-court inequality (2) becomes $[R-nY]/[\delta+\lambda-\lambda p] \ge R/(\delta+\lambda)$, which is algebraically equivalent to $p\lambda/(\delta+\lambda) \ge nY/R$. So with $Y=c(\delta+\lambda)/p$, we get $n \le Rp\lambda/[(\delta+\lambda)Y] = Rp^2\lambda/[c(\delta+\lambda)^2]$.

Notice $p = n^s/(n^s+m^s)$ implies $n = m[p/(1-p)]^{1/s}$. So with $n>0$ and $p>0$, we get

$$R\lambda p^{2-1/s}(1-p)^{1/s}/[c(\delta+\lambda)^2\ ] \ge m.$$

With $s>0.5$, the left side of this inequality is maximized by the probability $p = (2s-1)/(2s)$, and substituting this value of p yields the formula for M_0 in the proposition.

Proposition 3. Suppose that $s \ge 2/3$. If a force n is feasible against m for a leader with a weak court and $0 < n \le m$ then $w'(n|m) > 0$. So if m is globally feasible for leaders with weak courts then $\operatorname{argmax}_{k\ge 0} w(k|m) > m$.

Proof. The derivative with respect to n of the probability $p = p(n|m) = n^s/(n^s+m^s)$ is

$$p' = p'(n|m) = sm^sn^{s-1}/(n^s+m^s) = (1-p)ps/n.$$

The leader's pre-battle expected payoff is $w = w(n|m) = [Rp-nc(\delta+\lambda)]/[\delta+\lambda-\lambda p]$, and its derivative with respect to n satisfies

$$\begin{aligned}
w' &= [(p'R-c(\delta+\lambda))(\delta+\lambda-\lambda p) - (pR-nc(\delta+\lambda))(-\lambda p')]/(\delta+\lambda-\lambda p)^2 \\
&= [(R-nc\lambda)p' - c(\delta+\lambda-\lambda p)](\delta+\lambda)/(\delta+\lambda-\lambda p)^2 \\
&= [(R-nc\lambda)(1-p)ps - nc(\delta+\lambda-\lambda p)](\delta+\lambda)/[n(\delta+\lambda-\lambda p)^2\ ] \\
&= \{R(1-p)ps - nc[\delta+\lambda(1-p)(1+ps)]\}(\delta+\lambda)/[n(\delta+\lambda-\lambda p)^2\ ] \\
&\ge \{R(1-p)ps - [R\lambda p^2/(\delta+\lambda)^2][\delta+\lambda(1-p)(1+ps)]\}(\delta+\lambda)/[n(\delta+\lambda-\lambda p)^2\ ] \\
&= \{s(\delta+\lambda)^2 - \lambda p[\delta/(1-p)+\lambda(1+ps)]\}R(1-p)p/[n(\delta+\lambda)(\delta+\lambda-\lambda p)^2\ ] \\
&= \{s[\delta^2+2\lambda\delta+\lambda^2-\lambda^2p^2)] - \lambda[\delta p/(1-p)+\lambda p]\}R(1-p)p/[n(\delta+\lambda)(\delta+\lambda-\lambda p)^2\ ].
\end{aligned}$$

With $n\le m$, we have $p\le 0.5$, and so

$$\begin{aligned}
w' &\ge [s(\delta^2+2\lambda\delta+0.75\lambda^2) - \lambda(\delta+0.5\lambda)]Rp(1-p)/[n(\delta+\lambda)(\delta+\lambda-\lambda p)^2\ ] \\
&= [s(\delta+1.5\lambda) - \lambda](\delta+0.5\lambda)\ Rp(1-p)/[n(\delta+\lambda)(\delta+\lambda-\lambda p)^2\ ] > 0,
\end{aligned}$$

where the final strict inequality uses $s \ge 2/3 > \lambda/(\delta+1.5\lambda)$.

Now suppose that m is globally feasible with weak courts, so $v(m|m) \ge R/(\delta+\lambda)$. If some $n<m$ had $w(n|m) > w(m|m)$ then, with $p(n|m)<p(m|m)$, we would have

$$v(n|m) = w(n|m)/p(n|m) > w(m|m)/p(m|m) = v(m|m) \geq R/(\delta+\lambda),$$

and so n would be weak-court feasible against m. But then $w'(n|m)>0$, which implies that $n<m$ cannot maximize $w(n|m)$. That is, the maximum of $w(n|m)$ over $n\in[0,m]$ must be achieved at the top of the interval, at $n=m$. But $w'(m|m)>0$, and so some $k>m$ has $w(k|m) > w(m|m)$.

Proposition 4. When $s\leq 2$, the negotiation-proof equilibrium is $m_1 = Rs/[c(4\delta+2\lambda+s\lambda)]$. In this equilibrium, $m_1Y(m_1|m_1)/R = 2s(\delta+\lambda)/(4\delta+2\lambda+s\lambda)$. When $s\geq 0.763$, this equilibrium m_1 is greater than the bound M_0 from Proposition 2.

Proof Notice first that, with $p=p(n|m)$,

$$d/dn\,[n/p] = (p-np')/p^2 = [p-(1-p)ps]/p^2 = [1-(1-p)s]/p,$$

and so n/p is decreasing in n when $p<1-1/s$, but is increasing in n when $p\geq 1-1/s$. Then

$$w(n|m) = p(n|m)[R-c(\delta+\lambda)n/p(n|m)]/[\delta+\lambda-\lambda p(n|m)]$$

would be strictly increasing in n near any n where $p(n|m)<1-1/s$ and $w(n|m)>0$. Thus, for any m, the maximum of $w(n|m)$ must be achieved either at 0 or at some n such that $p(n|m)\geq 1-1/s$.

We have seen from the proof of Proposition 3 that,

$$w'(n|m) = [Rp(1-p)s - nc(\delta+\lambda(1-p)(1+ps))](\delta+\lambda)/[n(\delta+\lambda-\lambda p)^2]$$
$$= \{Rs - (n/p)c[\delta/(1-p)+\lambda(1+ps)]\}p(1-p)(\delta+\lambda)/[n(\delta+\lambda-\lambda p)^2]$$

So the sign of $w'(n|m)$ is determined by the expression $Rs - (n/p)c[\delta/(1-p)+\lambda(1+ps)]$. Over the set of all n such that $p\geq 1-1/s$, n/p and p are both increasing in n, and so this expression is decreasing in n and can cross 0 only once, at a value of n that has $w'(n|m)=0$ and so maximizes $w(n|m)$ in this set. Thus, if $n=0$ does not maximize $w(n|m)$, then the maximum of w must be achieved at the unique n such that $p(n|m)\geq 1-1/s$ and

$$Rs - (n/p)c[\delta/(1-p)+\lambda(1+ps)] = 0.$$

For m to be a negotiation-proof equilibrium, we need this equation to be satisfied with $n=m$, and we need $w(m|m) \geq w(0|m) = 0$. But $n=m$ implies $p(n|m)=0.5$, which satisfies $p\geq 1-1/s$ as long as $s\leq 2$. So the equilibrium conditions are

$$Rs = 2mc(2\delta+\lambda+\lambda s/2) \quad\text{and}\quad [R/2 - mc(\delta+\lambda)]/(\delta+\lambda/2) \geq 0.$$

The first condition is uniquely satisfied by m_1 in the proposition. The second condition holds iff

$$m \leq R/[2c(\delta+\lambda)].$$

With $s \le 2$, this condition is also satisfied at $m_1 = R/\{c[(4/s)\delta+(2/s+1)\lambda]\} \le R/[2c(\delta+\lambda)]$.

To show that m_1 is greater than M_0, we need

$$Rs/[c(4\delta+2\lambda+s\lambda)] > [R\lambda(2s-1)^{2-1/s}]/[4s^2c(\delta+\lambda)^2].$$

Letting $\theta=\delta/\lambda$, this inequality holds when

$$4(\theta+1)^2/(\theta+2+s) > (2s-1)^{2-1/s}/s^3$$

For any s, the left-hand side is minimized over $\theta \ge 0$ by letting $\theta=0$, and so the inequality holds if

$$4 > (2+s)(2s-1)^{2-1/s}/s^3,$$

which is true when $s \ge 0.76233$.

<u>Proposition 5.</u> When $1 < s \le 2$ and $\lambda \ge \delta(2-s)/(s-1)$, the oligarchic equilibrium is $m_2 = R[(s-1)\lambda-(2-s)\delta]/[c(\delta+\lambda)s\lambda]$. If $s<2$ then this oligarchic equilibrium satisfies $m_2 < m_1$, where m_1 is the negotiation-proof equilibrium for monarchs, but the ratio m_2/m_1 is increasing in λ/δ and s. If $s=2$ then $m_2=m_1$. On the other hand, if $s \le 1$ or $\lambda < \delta(2-s)/(s-1)$ then there is an oligarchic equilibrium at 0.

<u>Proof</u> The derivative of $\Omega(n|m) = -c +(R/n-\lambda c)p/(\delta+\lambda-\lambda p)$ with respect to n is

$$\Omega'(n|m) = \{(\delta+\lambda-\lambda p)[(R/n-\lambda c)p'-pR/n^2] + (R/n-\lambda c)p\lambda p'\}/(\delta+\lambda-\lambda p)^2$$
$$= [(\delta+\lambda)(R-n\lambda c)(1-p)ps - pR(\delta+\lambda-\lambda p)]/[n(\delta+\lambda-\lambda p)]^2$$
$$= \{R[\delta s-\delta/(1-p)+\lambda s-\lambda]-n(\delta+\lambda)s\lambda c\}(1-p)p/[n(\delta+\lambda-\lambda p)]^2.$$

So the sign of Ω' is determined by the expression

$$R[\delta s-\delta/(1-p)+\lambda s-\lambda] - n(\delta+\lambda)s\lambda c,$$

which is decreasing in n and so can cross 0 only once. Thus, if any $n>0$ maximizes $\Omega(n|m)$, then it must be at the unique solution of the equation

$$R[\delta s-\delta/(1-p)+\lambda s-\lambda] = n(\delta+\lambda)s\lambda c,$$

so that $\Omega'(n|m)=0$. For m to be an oligarchic equilibrium, this equation must be satisfied when $n=m$, but then $p = p(m|m) = 1/2$. So a oligarchic equilibrium must satisfy the equation

$$R[\delta s-2\delta+\lambda s-\lambda] = m(\delta+\lambda)s\lambda c,$$

which is uniquely satisfied by m_2 in the proposition. If $\lambda(s-1) > (2-s)\delta$, then this m_2 is positive and so is the unique oligarchic equilibrium.

The ratio of this m_2 over the monarchs' equilibrium $m_1 = Rs/[c(4\delta+2\lambda+s\lambda)]$ is

$$(4\delta+2\lambda+s\lambda)(s\delta-2\delta+s\lambda-\lambda)/[s^2(\delta+\lambda)\lambda] =$$
$$= [\lambda s+2(2\delta+\lambda)][s(\delta+\lambda)-(2\delta+\lambda)]/[s^2(\delta+\lambda)\lambda]$$
$$= [s^2\lambda(\delta+\lambda) - (2-s)(2\delta+\lambda)^2]/[s^2(\delta+\lambda)\lambda]$$
$$= 1 - (1+2\delta/\lambda)[1+\delta/(\delta+\lambda)](2/s-1)/s.$$

So when $s<2$, this ratio is less than 1, and it is increasing in λ/δ and s.

On the other hand, if $s\leq 1$ or $\lambda < \delta(2-s)/(s-1)$ then $(s-2)\delta+(s-1)\lambda < 0$, and so, for any n and m such that $n\geq m>0$, we get $p\geq 0.5$ and $[\delta s-\delta/(1-p)+\lambda s-\lambda] < 0$, and so $\Omega'(n|m) < 0$, as required for an oligarchic equilibrium at 0.

REFERENCES

Acemoglu, Daron, and James Robinson. 2006. Economic Origins of Dictatorship and Democracy Cambridge: Cambridge University Press.

Basu, Kaushik. 2000. Prelude to Political Economy. Oxford: Oxford U Press.

Baye, Michael R., Dan Kovenock, Casper G. de Vries. 1994. "The solutions to the Tullock rent-seeking game when R>2: mixed-strategy equilibria and mean dissipation rates." Public Choice 81(December):363-380.

Becker, Gary, and George Stigler. 1974. "Law enforcement, malfeasance, and compensation of enforcers." J Legal Studies 3(January):1-18.

Bol, Peter. 1998. "Government, society, and state: on the political visions of Ssu-ma Kuang and Wang An-shih." In Ordering the World: Approaches to State and Society in Sung Dynasty China, ed. by R. Hymes and C. Schirokauer, U. California Press, pp. 128-192.

Bueno de Mesquita, Bruce , Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. The Logic of Political Survival. MIT Press.

Egorov, Georgy, and Konstantin Sonin. 2005. "The killing game: reputation and knowledge in the politics of succession." CEPR discussion paper.

Egorov, Georgy, and Konstantin Sonin. 2006. "Dictators and their viziers: endogenizing the loyalty-competence trade-off." CEPR discussion paper.

Finer, Samuel E.. 1997. The History of Government from the Earliest Times. Oxford U Press.

Hardin, Russell. 1989. "Why a Constitution," In The Federalist Papers and the New Institutionalism, edited by B. Grofman and D. Wittman, NY: Agathon Press, pp. 100-

120.

Hume, David. 1748. "Of the Original Contract." Reprinted in Political Essays, edited by K. Haakonssen, Cambridge U. Press (1994), pp. 186-201.

Jouvenel, Bertrand de. 1948. On Power, translated by J. F. Huntington, Boston: Beacon Press.

Myerson, Roger. 2004. "Justice, institutions, and multiple equilibria." Chicago Journal of International Law 5(Summer):91-107.

Myerson, Roger. 2006. "Federalism and incentives for success of democracy." Quarterly Journal of Political Science 1(January):3-23.

Myerson, Roger. 2007. "Leadership, trust, and power." University of Chicago paper.

North, Douglass C. 1993. "Institutions and credible commitment." Journal of Institutional and Theoretical Economics 149(March):11-23.

Ray, Debraj, and Parikshit Ghosh. 1996. "Cooperation in community interaction without information flows," Review of Economic Studies 63(July):491-519.

Root, Hilton. 1989. "Tying the king's hands." Rationality and Society 1(October):240-258.

Schelling, Thomas C. 1960. The Strategy of Conflict. Harvard U. Press.

Shepsle, Kenneth. 1991. "Discretion, institutions, and the problem of government commitment." In Social Theory for a Changing Society, ed. by P. Bordieu and J. Coleman, New York: Russell Sage, pp. 245-263.

Skaperdas, Stergios. 1990. "Contest success functions." Economic Theory 7(June):283-290.

Svolik, Milan. 2006. "A theory of government dynamics in authoritarian regimes." University of Illinois working paper.

Tacitus, Cornelius. 1970. "The Germania." In The Agricola and the Germania, translated by H. Mattingly, London: Penguin Books.

Warren, W. L. 1973. Henry II. U. California Press.

Xenophon. 2001. The Education of Cyrus, translated by Wayne Ambler, Cornell U Press.

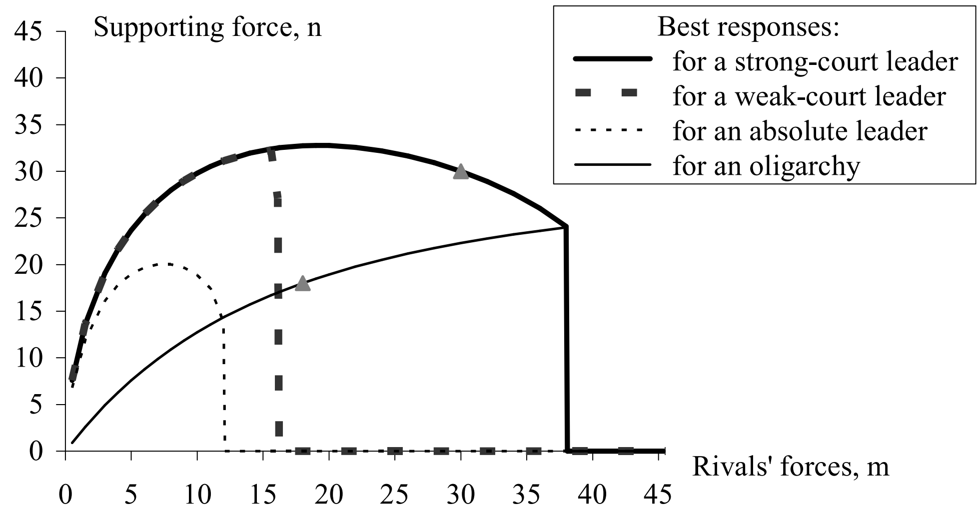

Figure 1. Optimal sizes of supporting forces for different kinds of regimes, depending on the anticipated rival forces, for an example with R=90, δ=0.05, λ=0.2, c=5, and s=1.5.

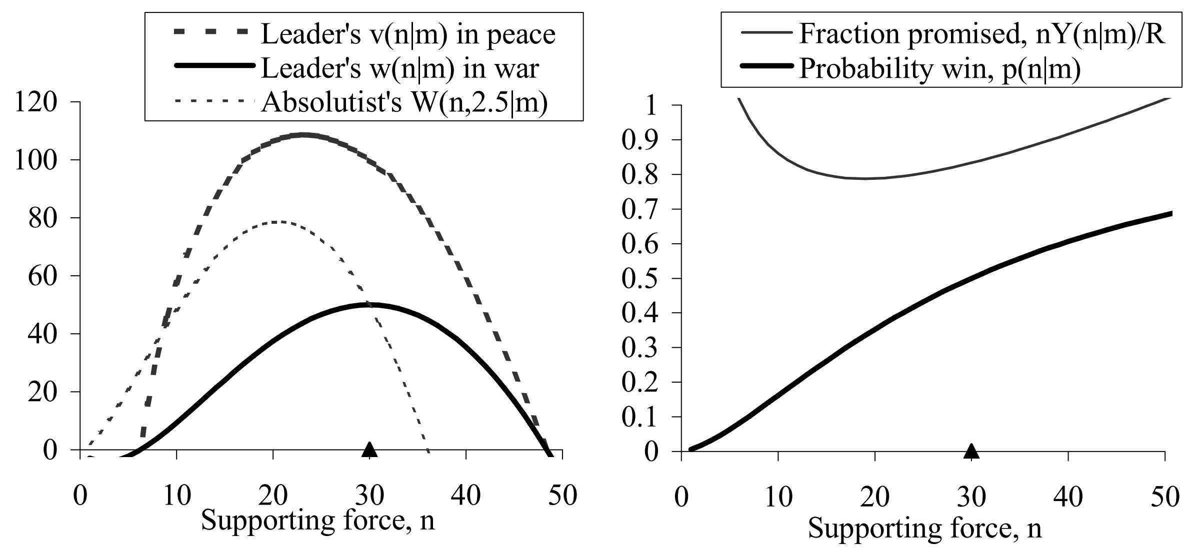

Figure 2. Results of changing the size of the supporting force, against m=30, with R=90, δ=0.05, λ=0.2, c=5, and s=1.5.

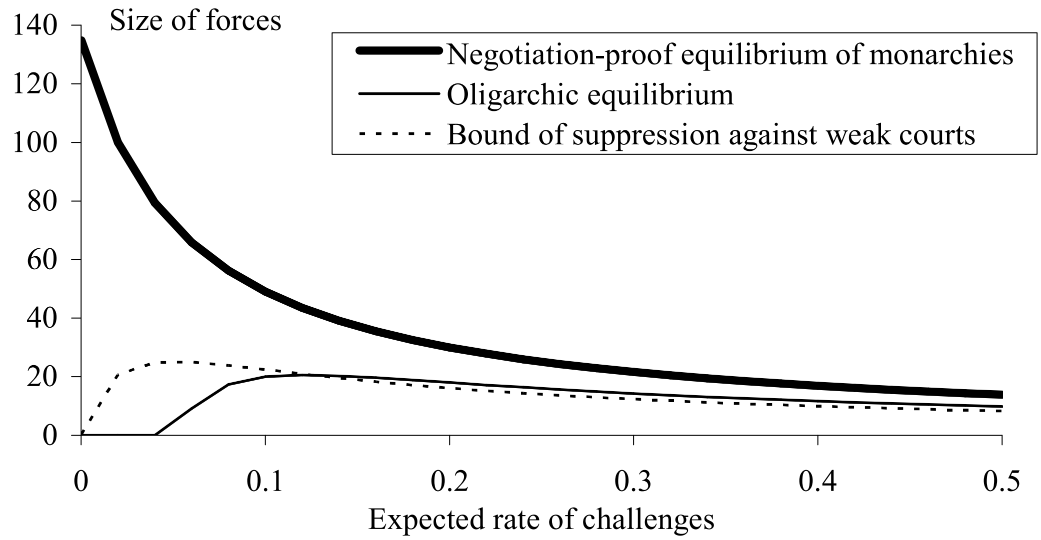

Figure 3. Changes in equilibrium for different values of λ, the expected rate of challenges, with R=90, δ=0.05, c=5, and s=1.5.

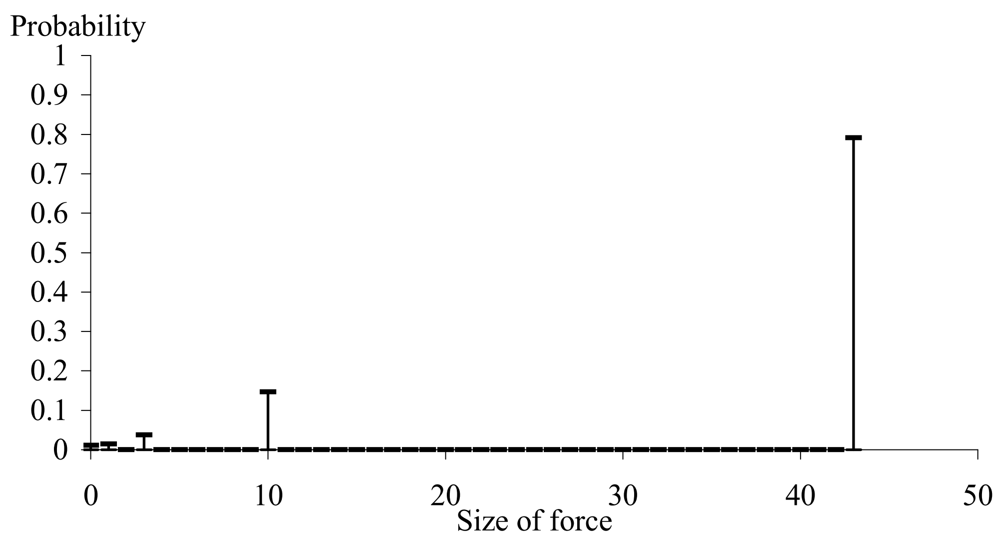

Figure 4. Discrete (integer) approximation of the randomized negotiation-proof equilibrium, for an example with R=90, δ=0.05, λ=0.2, c=5, and s=3.