A Theory of Leadership Based on Assignment of Information

Mana Komai and Mark Stegeman*

March, 2004 (revision: December, 2004)

Abstract

An organization makes collective decisions through neither markets nor contracts. Instead, rational agents voluntarily choose to follow a leader. In many cases incentive and coordination problems are solved: the unique nondegenerate equilibrium achieves the first best, even though every agent has incentives to free ride. The leader has no special talents but is distinguished by getting exclusive access to information. A crucial feature is that the leader reveals part but not all of her information. If agents have heterogeneous costs, then the optimal leader is often especially lazy (i.e., has high costs) and is never especially energetic.

1 Introduction

Researchers in management and political science believe that leadership is important, but economists rarely study it. In economists' models of organizations, managers mainly fill gaps, acting as owners' agents or exercising residual authority over subordinates when contracts are incomplete; politicians mainly represent amalgamations of voters' preferences. These models have produced important insights but omit crucial aspects of leadership. Leadership often includes encouraging subordinates and motivating by example.

This paper describes an "organization" which makes collective decisions through neither markets, nor contracts, nor voting, nor any grant of authority. Instead, rational agents choose to follow a leader absent any obligation to do so. In equilibrium, this strategy alleviates incentive as well as coordination problems.

The model has contrarian implications. Many researchers focus on reducing transactions costs by correcting information failures, but in our model efficiency requires

*We are grateful to seminar participants at Arizona, VPI, and the NBER/NSF Decentralization Conference, and especially Martin Dufwenberg, Stephen Gilliland, Mark Johnson, Zvika Neeman, and Roy Radner for comments and suggestions that improved the paper. The usual disclaimer applies.

an information failure. A huge literature studies how to align agents' goals through compensation, but our model abandons any attempt to align goals. Instead, only the leader gets access to critical information, and ignorance prevents her followers from *reaching* their goals. Collectively, however, the followers benefit from their ignorance, and they understand this.[1]

The model also offers a contrarian view of leaders themselves. They have no special talents and are distinguished merely by occupying the leadership position. Moreover, if agents are differentiated, then the optimal leader is either average or lazier than average. It is never optimal to appoint an unusually energetic leader.

This apparently gloomy picture of sluggish leaders and ignorant followers leads to happy results. In many cases incentive and coordination problems are completely solved, and the unique nondegenerate equilibrium achieves the first best – even though every agent has incentives to free ride.

All of these conclusions are, of course, the product of a simple model. In reality, organizations and leadership are complex phenomena, and this paper describes at best one small piece of the puzzle. Nonetheless, it is a piece that seems consistent with some commonplace observations.

In the sparse formal theory of leadership in economics, the nearest antecedent is Hermalin's (1998) model of a leader who has superior information. Hermalin's leader, however, fully reveals her information in equilibrium, and the resulting efficiency gains are qualitatively smaller than those obtaining here.

Sections 2 through 5 describe the basic results and Section 6 extends the analysis to heterogeneous agents. Section 7 discusses related literature, and Section 8 offers concluding remarks.

2 The Model: Actions and Payoffs

Each of $m > 1$ identical players, $i \in I = \{1, 2, ..., m\}$, must decide whether to join a proposed project. The main economic properties of the project are increasing but uncertain returns to participation, coupled with possibilities for free riding. Examples of such projects could include adopting new procedures or software that improve a firm's efficiency, helping to prepare a bid, or cooperating with a restructuring plan. In

[1] It is well known that information failures can increase efficiency. Section 7 discusses several received examples.

the political context, "projects" could include supporting controversial legislation or working for a candidate.

Formally, each player chooses an action from the set $A = \{N, P\}$, where N denotes non-participation and P denotes participation in the project. If all players choose N then each earns a payoff of zero. If all choose P then each earns a random payoff x, which is the same for all players and distributed on the interval $X = (\underline{x}, \overline{x})$ according to a strictly positive density $\phi(x)$. Assume $0 \in X$. A project is *good* if $x > 0$ or *bad* if $x < 0$.

If the players fail to coordinate, some choosing N and others choosing P, then player i's payoff is $\pi(a_i, q; x)$, where $a_i \in A$ denotes player i's action, q denotes the fraction of players who participate, and $\pi : A \times [0,1] \times X \rightarrow \Re_+$ is twice-continuously differentiable. (It is convenient to extend the domain of π to $q \in [0,1]$ rather than just multiples of $1/m$.) At the extremes of $q = 0$ and $q = 1$:

$$\pi(N, 0; x) = 0 \tag{1a}$$

$$\pi(P, 1; x) = x \tag{1b}$$

For q between zero and one, π must satisfy other assumptions, which, except as indicated, apply throughout its domain. Assumptions (1c) and (1d) imply that higher project quality (i.e., higher x) helps participants more than it helps nonparticipants and increases the (positive or negative) benefits that participants get from others' participation.[2]

$$\frac{\partial \pi}{\partial x}(P, q; x) > \max\left(0, \frac{\partial \pi}{\partial x}(N, q; x)\right) \quad for\ q > 0 \tag{1c}$$

$$\frac{\partial^2 \pi}{\partial q \partial x}(P, q; x) \geq 0 \tag{1d}$$

Assumption (1e) states that higher participation increases the (positive or negative) return to participation. This could reflect technical returns to scale or network effects, or social (e.g., "safety in numbers") effects. Assumption (1e) will simplify the analysis by eliminating equilibria in which some but not all players participate. Assumption (1f) similarly requires increasing returns to scale, but at the level of the group. The numerator of the derivative describes total surplus, and (1f) ensures that surplus-

[2] If participation means (for example) voting for an unpopular bill, then the political advantage of opposing the bill might increase as the quality of the bill decreases. In that case, $\frac{\partial \pi}{\partial x}(N, q; x) < 0$.

maximization requires either zero participation or full participation.[3]

$$\frac{\partial\left[\pi(P,q+\frac{1}{m};x)-\pi(N,q;x)\right]}{\partial q} > 0 \tag{1e}$$

$$\frac{\partial^2\left[qm\pi(P,q;x)+(1-q)m\pi(N,q;x)\right]}{\partial q^2} > 0 \tag{1f}$$

Assumptions (1g)-(1i) bound the circumstances that support participation. Given the previous assumptions, (1g) and (1h) imply that a player is willing to participate only if she learns something favorable about the realization of x *and* believes that others may also participate. Assumption (1i) states that a player does prefer to participate if she believes that x takes its maximal value and that everyone else will participate.

$$E_x\left[\pi(P,1;x)-\pi(N,1-\frac{1}{m};x)\right] < 0 \tag{1g}$$

$$\pi(P,\frac{1}{m};x) < 0 \tag{1g}$$

$$\pi(P,1;\overline{x})-\pi(N,1-\frac{1}{m};\overline{x}) > 0 \tag{1h}$$

Finally, assumption (1j) introduces an externality. It says that if one player defects from full participation, for a good or neutral (i.e., $x=0$) project, then she earns a positive payoff. In other words, the defector can take a free ride, gaining something from others' participation while sacrificing nothing herself. For bad projects, participation may (or may not) hurt nonparticipants, but assumption (1k) places a bound on how negative nonparticipants' payoffs can be. The purpose of this bound is only to rule out certain implausible equilibria; if one is willing to exclude those equilibria a priori, then (1k) is dispensable.[4]

$$\pi(N,1-\frac{1}{m};x) > 0 \qquad for\ x\geq 0 \tag{1j}$$

$$\pi(N,1-\frac{1}{m};x) > \pi(P,\frac{1}{m};x) \qquad for\ x<0 \tag{1k}$$

Example: Consider an organization with 7 identical players ($m=7$). Assume that project quality x is distributed uniformly on the interval $X=(-1,1)$, participants' payoffs are $\pi(P,q;x)=(x+1)q-1$, and nonparticipants' payoffs are $\pi(N,q;x)=$

[3] In assumption (1e), q represents the fraction of participants if a given player does not participate, $q+\frac{1}{m}$ represents the fraction of participants if he does participate, and the term in brackets shows how participation would change his payoff.

[4] In the proof of Lemma $2'$, (1k) rules out equilibria in which followers always do the opposite of what the leader does and the leader participates only for sufficiently negative x.

$(3+4x)q/12$. This setting satisfies assumptions (1a)-(1k). It is natural to interpret the constant term in $\pi(P,q;x)$ as a fixed participation cost.

The incentive to free ride can appear in unconventional ways. Suppose that the players are the Republican members of Congress, whose preferences are identical, and participation means voting for a bill. If all members vote for the bill, then the public perceives neither benefit nor harm and there is likewise no impact on the members' payoffs, so the quality of the bill is $x = 0$. In this case, assumption (1j) implies that any one member would benefit by defecting to vote against the bill. This benefit might arise because the defection calls attention to flaws in the bill and the defector gains a political reward $\pi(N, 1-\frac{1}{m}; 0) > 0$ for opposing a bill that his own action made less popular. (The realism of this scenario rests in the observation that majorities often place great value on getting a unanimous vote.) A more conventional interpretation of the same payoff structure is that the members like the content of the bill but their constitutents dislike it regardless of the vote breakdown, reducing the members' net benefits to zero even if they all support it. In that case, defecting and voting with the minority against the bill represents a familiar kind of free riding.

The payoff π may include nontransferable or noncontractible payoffs, as well as payoffs reflecting exogenous (explicit or implicit) contracts among the players. By taking π to be exogenous, we leave issues of contract design in the background.

That completes the discussion of the payoff function. The next several sections consider alternative specifications of the timing and information structure.

3 Discovering x.

If the players know nothing about the value of x when they make their participation decisions, then (1e) and (1g) imply that action N is strictly dominant for every player. If x is large, however, then it is individually and collectively rational for everyone to play P. This gives the players a reason to invest in discovering the value of x. Suppose that this is prohibitively expensive for individual players but worthwhile if players split the cost. Then discovering x is, in effect, a prior project, which also exhibits increasing returns and the possibility of free riding.

Suppose that each player can commit himself to paying k/m for research, where k denotes the total cost of discovering x, if every other player does the same. (In some contexts the collective commitment might represent the creation of an organization.) If any player fails to make this commitment, then the players pay nothing, learn nothing about x, and consequently do not participate in the project of Section 2. The premise of the sequel is that the players do make these commitments and discover x, and then they enter a new game with the payoff function π. If their expected equilibrium payoffs in that new game are high enough, then the initial research expenditure is individually rational.

This description of the research game is informal, because it serves merely to set the stage for the game of Section 2 and to motivate the main question of this paper: what happens after the players discover x?[5]

4 Model 1) Complete Information

The most obvious use of the players' information about x is to give it to everyone. Accordingly, consider the game in which all players observe x and then decide simultaneously whether to participate. Each player's strategy takes the form $s : X \to A$. Threshold strategies are natural: if a player adopts the *threshold strategy* t, then she plays P if and only if $x > t$.

Theorem 1 characterizes the (many) symmetric Nash equilibria in threshold strategies. All proofs are in the appendix.

Theorem 1 *There exists $\tau^C \in (0, \overline{x})$ such that symmetric adoption of the threshold strategy t constitutes a Nash equilibrium if and only if $t \geq \tau^C$.*

The threshold τ^C represents the lowest value of x at which each player is willing to participate if all others participate. If $x \in [\tau^C, \overline{x}]$, then the resulting subgame is a coordination game for which $q = 0$ and $q = 1$ are both Nash equilibria; if $x < \tau^C$, then action N is strictly dominant in the subgame.

The equilibria described by Theorem 1 have two unattractive features, which provide the motivation for this paper. First, the continuum of threshold equilibria may make it difficult to coordinate on a common threshold t. Coordinating on thresholds near τ^C may be especially difficult, because participation is weakly dominated at $x = \tau^C$ and "almost" weakly dominated when x exceeds τ^C only slightly. Therefore,

[5] The fundamental difference between the project of Section 2 and the prior research project described informally here is that participation in the prior project is collectively rational (and contractible) without acquiring any additional information.

players who are uncertain about others' strategies are likely to adopt thresholds significantly higher than τ^C. Such coordination failures imply missed opportunities for mutually beneficial and individually rational cooperation.

Second, even if the players somehow coordinate on the most efficient equilibrium, meaning that they coordinate on the threshold τ^C, they inefficiently fail to participate in a good project whenever $x \in (0, \tau^C)$. This inefficiency is the typical consequence of free riding.

Summarizing, the complete information game exhibits failures of both coordination and cooperation. The example of Section 2 illustrates both obstacles to efficiency.

Example (continued): The benefit to participation if all others participate is $\pi(P, 1; x) - \pi(N, 1 - \frac{1}{m}; x) = (10x - 3)/4$. Therefore, the lowest value of x that supports full participation is $\tau^C = 0.3$. This means that 70% of good projects are supportable as Nash equilibria, but 30% are undone by free riding. Moreover, if a player is uncertain of others' play, and he expects each of them to play N with independent probability of (for example) at least 20%, then it is irrational to participate for $x < 0.5$. Coordination problems can thus cause significant additional efficiency losses.

The next section shows how these cooperation and coordination problems can be solved by appointing a leader who has exclusive access to the value of x.[6]

5 Model 2) One Leader

A different way for the players to use their information about x is to let only one player see it. The players could choose such a "leader" as part of the agreement described informally in Section 3, granting her unique access to the results of the research.[7] Alternatively, they could pay the designated leader, to compensate her for becoming an expert and acquiring unique knowledge about the value of x. Whatever the interpretation, and regardless of whether the barriers to information flow are natural or artificially constructed, the point of this section is that restricting access to the value of x improves efficiency.

[6] If, given $x > 0$, the players could sign a cooperative agreement requiring all to play P if and only if they all sign, then the cooperation and coordination problems would also be solved. The premise of problems exhibiting moral hazard is that monitoring or other transactions costs, or nontransferable payoffs, make such a contract impractical.

[7] In equilibrium, the leader and followers will earn the same utility, so no player has an incentive to seek or to avoid the leadership position.

Formally, assume that exactly one player $l \in I$ observes x and that this *leader* acts first, choosing $a_l \in A$. Then the remaining *followers,* $I \backslash \{l\}$, observe the leader's action and act simultaneously. The leader's strategy takes the form $s_l : X \to A$ and each follower's strategy takes the form $s_f : A \to A$.[8] Because A has two elements, each follower has four possible strategies, $S_f = \{PP, NN, M, R\}$; PP means that the follower always participates, NN means never participate, M means that the follower mimics the leader's action, and R means reject the leader's action.

An *equilibrium* of this leader-follower game is a strategy profile $S^* = (s_l^*; s_f^*, f \in I \backslash \{l\})$ such that: s_l^* maximizes leader l's expected payoff given the followers' strategies and any $x \in X$; and s_f^* maximizes follower f's expected payoff given the other players' strategies.[9] Assumptions (1g) and (1h) imply that there always exists a trivial equilibrium in which no one ever participates.[10] Alternatively, an equilibrium is *productive* if someone participates with positive probability. The next lemma shows that any productive equilibrium must take a particular form: all followers must mimic the leader.[11]

Lemma 2 *If $(s_l^*; s_f^*, f \in I \backslash \{l\})$ is a productive equilibrium, then $s_f^* = M$ for all $f \in I \backslash \{l\}$.*

To aid intuition, this paragraph sketches the proof. If the leader never participates, then the followers get no information about x and therefore do not participate. Suppose that the leader participates with positive probability. Then all followers should adopt

[8] The formulation of the leader's strategy set does not allow her to reveal x directly, but in productive equilibria she would anyway have no incentive to do so (cf. Theorem 4): disclosing $x < 0$ would merely confirm that the equilibrium action N is strictly dominant, disclosing $0 < x < \tau^C$ would destroy beneficial cooperation, and disclosing $x > \tau^C$ could not improve upon the equilibrium outcome, which already achieves the first-best from the leader's viewpoint. In many realistic settings verifiable disclosure of x would impose significant costs of assembling and interpreting evidence, and other costs. For example, the photographs famously released during the Cuban missile crisis also revealed secrets about U.S. intelligence-gathering.

[9] For a "Bayesian" equilibrium, Harsanyi (1967) requires the leader to act optimally in a subset of X having probability one, which is equivalent to removing the reference to "any $x \in X$." That would reduce the present equilibrium to a standard Nash equilibrium, consistent with Harsanyi's demonstration that the set of Bayesian equilibria corresponds exactly to the Nash equilibria of the strategic form such that strategies are expressed as a function of type. For a "Bayesian Nash" equilibrium, Crawford and Sobel (1982) and sequels require that the sender act optimally given *any* realization of his type, analogous to requiring "any $x \in X$." The equilibrium concept employed here is thus a Bayesian Nash equilibrium in the sense of Crawford and Sobel rather than in the sense of Harsanyi.

[10] A no-participation equilibrium requires sufficiently many followers to play NN, with the remainder (if any) playing M.

[11] Note that equilibrium is defined for pure strategies only. Any mixing by the followers would be unstable in the sense that is familiar for symmetric coordination games: best replies to any small deviation in the mixture lead to a pure strategy equilibrium.

the same strategy, because the equilibrium participation of one follower implies, by increasing returns to participation (assumption (1e)), that other followers should also participate. That leaves four possibilities. Followers should not play PP, because the ex ante returns are too low (1g). Followers could play NN, but then the leader also should not participate (1h). If the followers play R, then the only reason for the leader to participate is to prevent followers from participating in very bad projects, but (1k) ensures that the followers cannot hurt the leader enough to cause her to participate in such unintuitive circumstances. The only remaining possibility is for all followers to mimic the leader (strategy M). The leader thus leads by example.

Lemma 2 implies that, in any productive equilibrium, the leader earns $\pi(P,1;x) = x$ if she participates or $\pi(N,0;x) = 0$ if she does not. Therefore, the leader adopts the threshold strategy $t = 0$.

In some cases, however, a productive equilibrium does not exist. The problem is that the leader's participation causes the followers to infer that $x > 0$, but N may be dominant for the followers given that event. In that case, the leader is not *credible*.[12] Intuitively, the followers expect the leader to lead them into too many projects that they would individually prefer to avoid, and they respond by ignoring the leader's signal. The next lemma characterizes the lowest participation threshold that the leader can adopt and leave the followers willing to follow.

Lemma 3 *There exists unique* $\tau^F \in (\underline{x}, \tau^C)$ *such that* $E_x[x - \pi(N, \frac{m-1}{m}; x) \mid x > \tau^F] = 0$.

Because she adopts the threshold $t = 0$ in any productive equilibrium, the leader is credible if and only if $\tau^F \leq 0$.

Lemma 3 implies that credibility is more likely when the benefits of free riding, as measured by $\pi(N, \frac{m-1}{m}; x)$, are relatively small. This in turn suggests that (all else equal) leaders of small organizations are more likely to be credible. Intuitively, the leader of a large organization acquires so much leverage that the discrepancy between her incentives and the incentives of the individual follower becomes too great to sustain credibility. Leaders' loss of credibility may thus be a previously underappreciated diseconomy of scale. Theorem 4 summarizes these findings.

[12] It is common to use the word "credible" to describe a threat that is believable. Here, a player is "credible" if she can send a signal that is informative in equilibrium.

Theorem 4 *There exists a productive equilibrium if and only if* $\tau^F \leq 0$. *In that case, the unique productive equilibrium is: the leader adopts the threshold strategy* $t = 0$ *and every follower adopts strategy* M.[13]

If the leader is credible, then Theorem 4 implies that the unique productive equilibrium of the leader-follower game produces full participation for strictly more values of x than do any of the complete information equilibria of Theorem 1. Moreover, a plausible application of forward induction discards the trivial non-productive equilibria: players who agree to fund research in the prior stage described in Section 3 presumably do not plan to play to an equilibrium in which the completed research is always ignored. In this sense, the leader-follower game completely solves the coordination problem apparent in Theorem 1.

The leader-follower game produces more participation than the complete information game because the leader's exclusive access to x allows her to "dupe" the followers into participating for values of x at which they would be unwilling to participate if they were fully informed. Collectively, however, the followers benefit from being so duped. The intuition is that the leader acts, in effect, as a representative agent who directs everyone's behavior. If she participates, then she and every other player earn $\pi(P, 1; x) = x$; if she does not participate then she and every other player earn $\pi(N, 0; x) = 0$. Therefore, she participates if $x > 0$ and does not participate if $x < 0$ and achieves the unconstrained first best.

A formal statement of the last claim requires a welfare measure. Let

$$W(q; x) = q\pi(P, q; x) + (1 - q)\pi(N, q; x)$$

denote the per capita surplus produced by the organization.

Theorem 5 *If* $\tau^F \leq 0$, *then the unique productive equilibrium of Theorem 4 achieves the first best (i.e., maximizes* $W(q; x)$ *over* q, *given any* x*).*

Leadership, artificially created by restricting access to information, thus solves both incentive conflicts and coordination problems.

[13] The definition of a threshold strategy arbitrarily requires that an indifferent leader not participate. The statement of Theorem 4 ignores, for simplicity, the alternative but essentially equivalent equilibrium resulting from modifying the leader's strategy so that she participates when $x = 0$. A similar caveat applies to Theorem 7.

Example (continued): If a leader is appointed, then the followers' "credibility threshold" is $\tau^F = -0.4$. Because $\tau^F \leq 0$, the leader is credible, she participates for all $x > 0$, and everyone else follows her.

Theorem 6 provides an unorthodox answer to the old question of the optimal degree of decentralization. Here, the efficient setting is at once completely decentralized and completely centralized. It is decentralized in the sense that there is no authority relationship. Each agent acts independently, without fear of sanction, earning payoffs according to the exogenous function π. Yet, in equilibrium, the leader effectively makes every decision.

6 Model 3) One Leader of a Heterogeneous Population

If $\tau^F > 0$, then the leader-follower game loses its efficiency in extreme fashion: a leader who is not credible cannot persuade anyone to participate and never participates herself. If players are heterogeneous, however, then the richer set of potential leaders improves the chance of finding one who is credible, even if she cannot achieve the first best. The choice of leader then becomes a substantial question.

Specifically, suppose that players vary according to their costs of participation. Up to this point participation costs have been implicit in the payoff function π, but now assume that each player $i \in I$ has an additional participation cost $c_i \equiv c + i\delta$ for some fixed $c, \delta > 0$. Player i's payoff is $\pi(P, q; x) - c_i$ if he participates or $\pi(N, q; x)$ if he does not. Assume that (1g)-(1i) and (1k) still hold for every player i, with $\pi(P, q; x) - c_i$ replacing $\pi(P, q; x)$ throughout. (The normalization (1b) holds as originally formulated.)

Participation costs have various interpretations. Player m could be the laziest player or the busiest, according to whether c_m represents a direct cost or an opportunity cost.

The following formula revises the welfare measure $W(q; x)$ to account for the participation costs c_i. It computes the highest surplus attainable if qm players participate, that is, it assumes that the participants are players $1, 2, ...qm$.

$$W_\delta(q; x) \equiv q\pi(P, q; x) + (1 - q)\pi(N, q; x) - q(c + [qm + 1]\delta/2)$$

The rest of the model, including the definition of equilibrium, is identical to Model 2.

Leaders and followers with different participation costs obviously differ in their willingness to participate. For each follower f, Lemma 6 defines the threshold τ_f^F, analogous to τ^F in Lemma 4 For each potential leader $l \in I$, Lemma 6 likewise defines the threshold τ_l^L at which leader l is indifferent to participating if every follower mimics her. (In Model 2, that threshold was zero for any leader.) Lemma 6 makes the elementary observation that players with higher participation costs are less willing to participate, both as leaders and as followers.

Lemma 6 *For each $l \in I$, there exists unique $\tau_l^L \in (0, \overline{x})$ such that $\pi(P, 1; x) - c_l = 0$, and τ_l^L is strictly increasing in l. For each $f \in I$, there exists unique $\tau_f^F \in X$ such that $E_x[\pi(P, 1; x) - \pi(N, \frac{m-1}{m}; x) \mid x > \tau_f^F] - c_f = 0$, and τ_f^F is strictly increasing in f.*

Heterogeneous participation costs raise the possibility of equilibria in which only some followers participate, but the results in this section suppress such (complicated) outcomes by assuming that δ is small enough to ensure that all followers adopt strategy M in any productive equilibrium. Then the credibility constraint for a productive equilibrium, which was $\tau^F \leq 0$ in the model of identical players, becomes $\tau^F_{\max(I \setminus \{l\})} \leq \tau_l^L$. If this inequality holds for a given leader l, then the signal conveyed by leader l's participation is strong enough to support participation by everyone else. Leaders with higher participation costs are more likely to be credible, because (in equilibrium) followers know that they are more selective when deciding which projects are worthwhile.

Theorem 7 describes equilibria similar to those of Theorem 4. For any given leader, Theorem 7 shows that the uniqueness of the productive equilibrium survives small differences in participation costs.

Theorem 7 *If $\delta > 0$ is sufficiently small given m and π, then the following is true. For any $l \in I$, there exists a productive equilibrium if and only if $\tau^F_{\max(I \setminus \{l\})} \leq \tau_l^L$. In that case, the unique productive equilibrium is: the leader l adopts the threshold strategy $t = \tau_l^L$ and every follower adopts strategy M.*[14]

The question of efficiency is more complicated. Model 2 produces first-best outcomes because a credible leader's preferences mirror those of her followers, and the present model can similarly achieve the first-best if the (credible) leader is an "aver-

[14] The definition of a threshold strategy arbitrarily requires that an indifferent leader not participate (cf. fn. 13).

age" player. Let $a \equiv (m+1)/2$ and for the moment assume that m is odd; then player $a \in I$ is the "average" player.

Theorem 8 *If m is odd and $\delta > 0$ is sufficiently small given m and π, then the following is true. If $\tau_m^F \leq \tau_a^L$, then the unique productive equilibrium induced by leader $l = a$ achieves the first best (i.e., maximizes $W_\delta(q; x)$ over q, given any x).*

Suppose instead that $\tau_a^L < \tau_m^F$, so player a is not credible. Then a lower-cost player would also not be credible, because she would try to induce participation for even lower values of x. A higher-cost player might be credible, but she will not achieve the first best, because she is not representative and will skip some projects that would increase average surplus. Navigating between these inefficiencies, the best leader is whichever player has the lowest cost of participation, subject to the constraint of being credible.

To make this claim precise, let τ_l^* denote the participation threshold of any given leader $l \in I$, assuming that she and her followers coordinate on the productive equilibrium if one exists. Formally, $\tau_l^* = \tau_l^L$ if $\tau_{\max(I \setminus \{l\})}^F \leq \tau_l^L$, and $\tau_l^* = \overline{x}$ otherwise. By adopting threshold τ_l^*, leader l generates the following expected surplus:

$$\widehat{W}_\delta(l) \equiv \int_{\tau_l^*}^{\overline{x}} W_\delta(1; x)\phi(x)dx$$

If there is any player who would be credible as the leader, then Theorem 9 describes the optimal leader(s), meaning the one who maximizes $\widehat{W}_\delta(l)$.

Theorem 9 *Assume that $\tau_{\max(I \setminus \{l\})}^F \leq \tau_l^L$ for some $l \in I$, and let $\widetilde{l}$ denote the smallest such l. If $\delta > 0$ is sufficiently small given m and π, then the following is true. If $\widetilde{l} \geq a$, then $\widetilde{l}$ is the unique maximizer of $\widehat{W}_\delta(l)$. Otherwise, any maximizer of $\widehat{W}_\delta(l)$ is an element of $\{a - \frac{1}{2}, a, a + \frac{1}{2}\}$.*

Theorem 9 implies that it is never optimal to appoint a leader who has unusually low costs of participation, for example, a leader who is especially energetic or who has more time than most to commit to the job.

Example (continued): To the example of Section 2 add the participation costs $c = .16$ and $\delta = .01$, implying $c_1 = .17$, $c_2 = .18$, ..., $c_6 = .22$, and $c_7 = .23$. (Note that assumptions (1g)-(1i) and (1k) still hold.) Then $\tau_l^L = .16 + .01l$ for each potential leader $l \in I$, and $\tau_f^F = .048 + .028f$ for each potential follower $f \in I$. For the average leader ($l = a = 4$), $\tau_a^L = .20 < .24 = \tau_{\max(I\setminus\{l\})}^F$, violating the credibility condition. Leaders $l = 5$ and $l = 6$ are similarly not credible. If, however, the "laziest" player $l = 7$ is appointed to lead, then $\tau_l^L = .23 > .22 = \tau_{\max(I\setminus\{l\})}^F$, and she is credible. Promoting the laziest (or busiest) player to leadership has the double advantage of moving her from a position in which her "laziness" harms the organization's efficiency to a position in which it can enhance efficiency; in other words, it helps both sides of the inequality.

Because leader $l = 7$ participates for $x > .23$, while participation would be efficient for $x > .20$, she "wastes" 3% of all good projects. This is much better performance than accrues under full information, which given the additional participation costs c_i can support full participation only for $x > .62$.[15]

7 Related Literature

Hermalin (1998) studies a team leader who, like ours, has private information about the return to effort and increases observable effort when the return is high. His leader-follower equilibrium produces more efficient outcomes than does full information, because it improves the leader's incentives to work. Although we rule out the transfer schemes that Hermalin uses to enhance efficiency, our leader-follower equilibrium generates qualitatively more surplus, because (unlike in Hermalin's model) the leader does not fully reveal her information and this causes everyone to work harder, not just the leader. Vesterlund (2003) applies Hermalin's idea to a model of charitable contributions, but in her model the leader chooses whether to acquire information before deciding whether to contribute. Vesterlund focuses on whether a third player, a fundraiser who moves first, chooses to announce that the leader's contribution will be public. Andreoni (2004) builds on Vesterlund's (2003) model by endogenizing the selection of the leader. In Andreoni's model, agents with low costs of contribution (e.g., the rich) step forward to accept this costly role. Vesterlund's and Andreoni's leaders, like Hermalin's, fully reveal their information.[16]

[15] One may wonder whether a leader $l < 7$ could credibly lead a *subset* of relatively low cost followers, but it is easy to confirm that the returns to scale in this example are sufficient to imply that any leader $l < 7$ cannot credibly lead any number of followers. No player is ever willing to participate unless at least five others do so, and the credibility constraint fails by a wide margin for a leader with exactly five followers, because the followers are willing to participate only for the very highest values of x.

[16] These results stand in contrast to Varian's (1994) benchmark result that, due to crowding out, total contributions to a public good are higher if they occur simultaneously than if they occur sequentially. In a complete information environment, Romano and Yildirim (2001) show that warm-glow effects can

[Author(s)] (2002,2004) [published working papers] introduce the idea that partial revelation of a leader's information improves efficiency and can produce first best outcomes. Those papers extend the present analysis to a continuous action set and, separately, to the possibility of solving credibility problems by dividing authority among several leaders. [Author(s)] (2002) also study the case of the linear payoff function $\pi_i = 1 - a_i + (a_1 + a_2 + ...a_m)x$, where $a_i = 1$ for any participant and $a_i = 0$ for any nonparticipant; this payoff function differs from the present model because players who move simultaneously have dominant strategies, eliminating the coordination issue. The linear case lies on the cusp of discrete and continuous action spaces because the discrete actions can be interpreted as the endpoints of a compact choice set.[17]

A substantial literature describes potential advantages to ignorance. A classic example is the provision of insurance, but Ostrovsky and Schwarz (2004) extend that story to a setting in which all agents are risk neutral and total surplus is fixed. In their model, individual students among a large population may have an incentive to secure jobs long before graduation, to protect themselves against bad academic performance; a concave mapping from performance to job-related rents induces a demand for insurance. In a game among schools who seek to maximize their students' rents, the schools insure their students by withholding some information about performance. In another recent study, Levy and Razin (2004) describe a game between two countries deciding how much to cooperate. The key feature is that a democracy must share with its rival all information possessed by its decision-maker, the public. This constraint allows democracies to cooperate where autocracies cannot. (In contrast, our leader-follower model enhances cooperation because an informed player uses information that she *cannot* share with her rivals.) Levy and Razin show that in some cases democracies cooperate better if decision-makers in both countries are completely uninformed.[18]

---

overturn Varian's conclusion and create advantages to sequential giving.

[17] The first draft of [author(s)] (2002), which proposed and solved the linear case, was apparently the first recognition of the efficiency of partial revelation of a leader's information. We are grateful to Lise Vesterlund for her encouraging reply to that draft and for subsequently serving on [author's] dissertation committee.

Potters, Sefton, and Vesterlund (2004) test the linear case experimentally, for a two-player game, that is, with payoff function $\pi_i = 1 - a_i + (a_1 + a_2)x$ and $a_i \in \{0, 1\}$ (e.g. [author(s)] (2002) Example 2). They add a prior voting stage in which the two players must simultaneously affirm that the informed player will contribute first; otherwise the players contribute simultaneously. The experimental results broadly confirm the predictions of the model, in the sense of supporting the efficient equilibrium.

[18] Levy and Razin's model creates an explicit role for a leader. Each of their democracies has a leader who must decide whether to reveal information, while the public makes the decision. The leader and the public have common interests; the leader essentially decides, through cheap talk, whether the public should make the decision in a (completely) informed or (completely) uninformed way, mindful

Our model is also related to the idea of information cascades, but unlike the many studies emphasizing the inefficiencies induced by cascades, we use the leader-follower relationship to improve efficiency.

The management literature includes myriad studies of leadership, many of them empirical and most omitting the formal modeling familiar to economists. Such studies address what leaders do, how they do it, how they can do it better, how to adjust their environment so that they can do it better, and which personal attributes are important for leadership, in various settings.[19] Economists have made relatively few contributions to this literature, but Rotemberg and Saloner (1993), for example, present a model that compares the effectiveness of selfish and empathetic managers in different situations. The recent management literature (e.g., Case (1995)) tends to promote sharing information with employees, but this paper provides a reason to doubt that such policies are always wise.[20] Canice Prendergast (1993) presents a quite different model with a related message. She shows that if managers rely on information provided by workers, then workers' incentive to conform means that it may be best to insulate them from managers' other sources of information.

8 Summary and Concluding Remarks

The game of Section 4 exhibits the familiar problems of free riding and coordination failure, but our remedy reverses the usual method of mechanism design. Instead of using contracts to align agents' incentives, given an exogenous information structure, we leave contracts in the background and redesign agents' information. Instead of trying to improve information, subject to monitoring and processing costs, we keep critical data away from decision-makers. By depriving agents of the fine information required for profitable defections, our low-cost information structure promotes cooperation as well as coordination.

Organization theory often overlooks the collective benefits that subordinates may derive from their own ignorance. A similar point applies to voters.

This paper shows that it is possible to address important issues of organizational

that the rival will see the information also.

[19] Dansereau (1998) surveys some of these studies.

[20] Milgrom and Roberts (1992) survey general themes from both economists' formal models of organizations and the less formal management literature. They note the consensus view that the "key problem in achieving effective coordination and adaptation is that the information needed to determine the best use of resources... is not freely available to everyone."

design without invoking contracts, prices, residual authority, or bargaining.[21] Instead, the simple equilibrium rule of following the leader produces efficient behavior without monetary payments or ex post coercion. This minimalist theory of organizations yields a minimalist theory of leadership. Our leaders have no special skills or authority and yet are crucial for sustaining efficiency.

Our model is unusual in correcting a leader's incentives by magnifying her impact rather than her compensation. When efficiency fails, as might occur in a large organization, it is because the leader has too much impact – she becomes too pivotal – relative to the other players. To restore cooperation in such cases, a leader must find a way to convince her followers that she will not (literally) mislead them. We have shown that one way to restore the leader's credibility is to appoint a high-cost leader, because such a leader has a greater incentive to shirk. The discrete gain from restoring credibility dominates the loss from shirking at the margin.

Common complaints that unproductive persons rise to leadership positions may thus reflect rational selection. Indeed, folk wisdom suggests that there may be advantages to choosing leaders who are too lazy or too busy to lead (and not because busyness signals talent). Yet few if any prior models explain why this might be so.

Our simple model inevitably ignores many issues that are important for real organizations. In reality, leaders perform tasks beyond those modeled here, subordinates get information that is useful to leaders, agents play asymmetric roles in team production, leaders can build reputations that enhance their credibility, computation constraints bind, etc. (Which issues are most important depends on the setting; managing firms differs from political organizing.) All of these complications may affect our conclusions. Yet, while we study only one slice of the complex reality of organizations, we believe that it contains enough truth to contribute to our overall understanding of leadership.

To keep the focus on information (and reflecting our belief that utility may be less transferable than is often assumed), we have suppressed transfers and any differences in the payoff functions of leaders and followers. A more complete analysis would consider how various transfer schemes might affect our conclusions and, conversely, how the issues identified in this paper may affect the design of optimal contracts.

[21] The prior research stage described informally in Section 3 requires bargaining to a simple contract.

9 Appendix

Consider Model 3, generalized to allow $c, \delta \geq 0$. Models 1 and 2 appear as the special case $c = \delta = 0$ (with different strategy sets for Model 1) and Model 3 appears as the special case $c, \delta > 0$. It is useful to define summary notation for player i's gain from participation, as a function of q_{-i}, the fraction of the remaining players who participate. If player i participates, then $q_{-i} = (mq - 1)/(m - 1)$; if not, then $q_{-i} = mq/(m - 1)$. For $i \in I \cup \{a\}$ (recall $a \equiv \frac{m+1}{2}$), let:

$$\beta_i(q_{-i}; x) \equiv \pi\left(P, \frac{(m-1)\, q_{-i} + 1}{m}; x\right) - \pi\left(N, \frac{(m-1) q_{-i}}{m}; x\right) - c_i.$$ [22]

Assumption (1e), and assumptions (1c) and (1d), imply respectively:

$$\partial\beta_i/\partial q_{-i} > 0; \quad \partial\beta_i/\partial x > 0 \tag{A1}$$

Assumptions (1g) and (1i) (which for Model 3 hold for all c_i) can be restated as:

$$E_x\left[\beta_i(1; x)\right] \; < \; 0 \quad \text{for all } i \tag{1g$'$}$$

$$\beta_i(1; \overline{x}) \; > \; 0 \quad \text{for all } i \tag{1i$'$}$$

Given $\partial\beta_i/\partial q_{-i} > 0$, (1g$'$) implies

$$E_x\left[\beta_i(q_{-i}; x)\right] < 0 \quad \text{for all } i \tag{A2}$$

It is similarly useful to define notation for leaders' and followers' gain from participation, in equilibrium. For any $l \in I \cup \{a\}$ and $f \in I$, let

$$h_l(x) \; \equiv \; \pi(P, 1; x) - c_l$$

$$k_f(t) \; \equiv \; E_x[\beta_f(1; x) \mid x > t]$$

Lemma 0. For any $i, l, f \in I \cup \{a\}$:

(i) There exists unique $\tau_i^C \in (0, \overline{x})$ such that $\beta_i(1; \tau_i^C) = 0$.

(ii) $h_l' > 0$ and there exists unique $\tau_l^L \in [0, \tau_l^C)$ such that $h_l(\tau_l^L) = 0$.

(iii) $k_f' > 0$ and there exists unique $\tau_f^F \in (\underline{x}, \tau_f^C)$ such that $k_f(\tau_f^F) = 0$.

Proof. Assumptions (1b) and (1j) imply $\beta_i(1; 0) < 0$ and $h_l(0) \leq 0$ for all i,$l \in I$. Result (A1) and assumption (1c) imply $\partial\beta_i(1; x)/\partial x > 0$ and $h_l' > 0$ for all i,$l \in I$.

[22] Substituting for q_{-i} using the case-dependent equalities, one can write $\beta(q_{-i}; x) = \pi(P, q; x) - \pi(N, q; x)$, but in that case (unlike in assumptions (1e) and (1f)) the two values of "q" differ by $1/m$, according to whether player i participates.

Claim (i) follows from (1i$'$). By definition $\beta_l(1;\tau_l^C) = h_l(\tau_l^C) - \pi\left(N, \frac{m-1}{m}; \tau_l^C\right) = 0$, so (1j) implies $h_l(\tau_l^C) > 0$. Claim (ii) follows. Let $\Phi(x) = \int_0^x \phi(t)dt$. Using $\partial\beta_f/\partial x > 0$ (A1), we have, for $t \in (0, \overline{x})$:

$$\begin{aligned} k_f'(t) &= \frac{\partial E_x[\beta_f(1;x) \mid x > t]}{\partial t} \\ &= \frac{\phi(t)}{1-\Phi(t)} \times \left[\int_t^{\overline{x}} \frac{\beta_f(1;x)\phi(x)dx}{1-\Phi(t)} - \beta_f(1;t)\right] \\ &> \frac{\phi(t)}{1-\Phi(t)} \times \left[\int_t^{\overline{x}} \frac{\beta_f(1;t)\phi(x)dx}{1-\Phi(t)} - \beta_f(1;t)\right] = 0 \end{aligned}$$

Therefore, $k_f'(t) > 0$. From (1g$'$): $k_f(\underline{x}) = E_x[\beta_f(1;x)] < 0$. By definition, $\beta_f(1;\tau_f^C) = 0$, which with $\partial\beta_f/\partial x > 0$ (A1) and $\tau_f^C < \overline{x}$ implies $k_f(\tau_f^C) > 0$. These facts establish (iii) $\square$

Proof of Theorem 1. For any $i, j \in I$, $c_i = c_j$ implies $\beta_i = \beta_j$ and $\tau_i^C = \tau_j^C$ (from Lemma 0(i)). Let $\tau^C \in (0, \overline{x})$ and β denote these common values. Assumption (1h) implies: (a) given any x, it is a Nash equilibrium for no one to participate. If every player participates, then one player's gain from participation is $\beta(1, x)$. The definition of τ^C and (A1) imply that $\beta(1,x) < 0$ for $x < \tau^C$ and $\beta(1,x) > 0$ for $x > \tau^C$. Therefore: (b) full participation is a Nash equilibrium iff $x \geq \tau^C$. Statements (a) and (b) imply the result. $\square$

It is convenient to prove the remaining results in a sequence different from the sequence of appearance. Lemma 2$'$ generalizes Lemma 2 to the case of heterogeneous players. After proving results 6-9 for the heterogeneous case, results 2-5 for the homogeneous case are little more than special cases.

Lemma 2$'$. If $\delta > 0$ is sufficiently small given m and π, or $\delta = 0$, then the following is true. Any productive equilibrium $(s_l^*; s_f^*, f \in I\backslash\{l\})$ has $s_f^* = M$ for all $f \in I\backslash\{l\}$.

Proof. Because β_i is C^2, (A1) implies that $\partial\beta_i(q_{-i};x)/\partial q_{-i} > \omega$ for all q_{-i}, x, and i, for some fixed $\omega > 0$. Therefore, $\beta_i(q_{-i};x) - \beta_i(q_{-i} - \frac{1}{m-1};x) > \frac{\omega}{m-1}$, implying $\beta_i(q_{-i};x) - \beta_j(q_{-i} - \frac{1}{m-1};x) > \frac{\omega}{m-1} - \delta m$, for all q_{-i}, x, i, and j. Assume $\delta \in (0, \frac{\omega}{(m-1)m})$. Then

$$\beta_i\left(q_{-i};x\right) > \beta_j\left(q_{-i} - \frac{1}{m-1};x\right) \tag{A3}$$

To study a follower's optimization problem, define the following notation. For $a_l \in A$, let $X(a_l) \equiv \{x \in X \mid s_l^*(x) = a_l\}$, the inverse image of the leader's action, and for $f \in I \backslash \{l\}$ let $q_{-f}(a_l)$ denote the fraction of the remaining $m-1$ players who participate, as implied by $(s_f^*, f \in I \backslash \{l\})$. Let $\psi_f(a_l)$ denote f's gain from participation, given a_l:[23]

$$\psi_f(a_l) \equiv E_x[\beta_f(q_{-f}(a_l); x) \mid x \in X(a_l)]$$

The equilibrium condition for follower f requires that, for each a_l such that $X(a_l)$ occurs with positive probability:

$$\psi_f(a_l) > 0 \text{ implies } s_f^*(a_l) = P; \quad \psi_f(a_l) < 0 \text{ implies } s_f^*(a_l) = N$$

For the given equilibrium, the proof has five steps.

(i) No follower chooses strategy PP. Given q_{-i}, strategy PP earns $E_x\left[\beta_f(q_{-i}; x)\right]$ more than strategy NN, but (A2) implies that this is strictly negative. Therefore, PP is strictly dominated.

(ii) The probability that the leader participates is neither zero nor one. Suppose that $X(P)$ occurs with probability one. Then M is equivalent to PP and strictly dominated by (i), implying that every follower chooses R or NN, implying that every follower plays action N, but then (1a) and (1h) imply that the leader should never participate, a contradiction. Suppose that $X(N)$ occurs with probability one. Then (i) again implies that every follower plays action N, but that contradicts the assumption of a productive equilibrium.

(iii) All followers choose the same strategy. Fix $a_l \in A$. It is sufficient to show that $s_f^*(a_l)$ takes the same value for all $f \in I \backslash \{l\}$. Suppose not. Let $\left\{F^N, F^P\right\}$ denote the partition of $I \backslash \{l\}$ such that $s_f^*(a_l) = N$ for all $f \in F^N$ and $s_f^*(a_l) = P$ for all $f \in F^P$. Choose $f' \in F^P$ and $f \in F^N$. Then $q_{-f'}(a_l) = q_{-f}(a_l) - \frac{1}{m-1}$, implying from (A3) $\beta_f\left(q_{-f}(a_l); x\right) > \beta_{f'}(q_{-f'}(a_l); x)$ for all x, which implies $\psi_f(a_l) > \psi_{f'}(a_l)$, but, given (ii), this contradicts the premise that follower f' chooses P while f does not.

(iv) That strategy cannot be NN. If all followers choose NN, then (1a) and (1h) imply that the leader should never participate, contradicting (ii).

(v) That strategy cannot be R. If all followers choose R, then leader l's gain from participating given $x \in X(P)$ is $\pi(P, \frac{1}{m}; x) - \pi(N, 1 - \frac{1}{m}; x) - c_l \geq 0$, but that contradicts (1j) or (1k). $\square$

[23] The ψ_f and k_f functions differ because ψ_f describes a follower's gain from participating given an arbitrary action by the leader and the arbitrary strategy profile $(s_l^*; s_f^*, f \in I \backslash \{l\})$.

Proof of Lemma 6 (including the case $c = \delta = 0$)**.** Lemma 0 establishes the existence of $\tau_l^L \in [0, \overline{x})$. Assumption (1b) implies that $\tau_l^L = 0$ iff $c = \delta = 0$. For the rest of the proof, assume that $\tau_l^L \neq 0$. Consider $\widehat{l} > l$. Because $c_{\widehat{l}} > c_l$, the definition of h_l implies $h_{\widehat{l}}(\tau_l^L) < h_l(\tau_l^L)$. The definition of τ_l^L implies that $h_l(\tau_l^L) = 0 = h_{\widehat{l}}(\tau_{\widehat{l}}^L)$. Therefore, $h_{\widehat{l}}(\tau_l^L) < h_{\widehat{l}}(\tau_{\widehat{l}}^L)$, but Lemma 0(ii) shows that $h'_{\widehat{l}} > 0$, so $\tau_l^L < \tau_{\widehat{l}}^L$. Lemma 0(iii) similarly establishes the existence of τ_f^F and shows that τ_f^F is strictly increasing in f. □

Proof of Theorem 7 (including the case $c = \delta = 0$)**.** Assume that $\delta > 0$ is small in the sense required for Lemma 2′. Fix $l \in I$. Lemma 2′ shows that a productive equilibrium requires all followers $f \in I \backslash \{l\}$ to play M. Given this, it is sufficient to show:

(i) If all followers play M, then the unique optimal strategy for leader l is the threshold strategy τ_l^L.[24]

(ii) If the leader plays the threshold strategy τ_l^L, and all other followers play M, then M is optimal for follower f iff $\tau_l^L \geq \tau_f^F$.

Proof of (i): Leader l earns payoff $\pi(P, 1; x) - c_l$ from playing P or zero (given (1a)) from playing N. Therefore, using Lemma 0(ii), leader l's equilibrium condition reduces to: if $x > \tau_l^L$ then $s_l^*(x) = P$; if $x < \tau_l^L$ then $s_l^*(x) = N$.

Proof of (ii): Given other players' strategies, (1a), and (1h): $\psi_f(P) = k_f(\tau_l^L)$ and $\psi_f(N) < -c_f$, as these functions are defined for Lemma 0 and the proof of Lemma 2′. The statement of follower f's equilibrium condition in the proof of Lemma 2′ shows that M is optimal for follower f iff $\psi_f(P) \geq 0$ and $\psi_f(n) \leq 0$. Therefore, M is optimal iff $k_f(\tau_l^L) \geq 0$, which by Lemma 0(iii) is true iff $\tau_l^L \geq \tau_f^F$. □

Proof of Theorem 8 (including the case $c = \delta = 0$)**.** The first part of the proof does not require m odd. Let

$$B(q, x) \equiv \frac{\partial^2 \left[q\pi(P, q; x) + (1 - q)\pi(N, q; x) \right]}{\partial q^2}$$

Because π is C^2, (1f) implies that $B(q, x) > \omega$ for all $(q, x) \in [0, 1] \times X$, for some fixed $\omega > 0$. Therefore, $\partial^2 W_\delta(q; x)/\partial q^2 = B(q, x) - m\delta > \omega - m\delta$. Assume $\delta \in [0, \frac{w}{m})$ and that δ is small in the sense required for Theorem 7. Then $\partial^2 W_\delta(q; x)/\partial q^2 > 0$, which implies that, given any x, $W_\delta(q; x)$ is maximized at $q = 0$ or $q = 1$. Because (1a)

[24] This claim disregards the inconsequentially differentiated strategies such that the leader participates at the threshold $x = \tau_l^L$ (cf. fn. 13).

implies that $W_\delta(0;x) = 0$, the sign of $W_\delta(1;x) = \pi(P,1;x) - c - a\delta$ indicates whether $W(q;x)$ is maximized at $q = 0$ or $q = 1$.

The definition of τ_a^L (Lemma 0(ii)) implies $\pi(P,1;\tau_a^L) - c_a = 0$, implying $W_\delta(1;\tau_a^L) = 0$. Because (1c) implies $\partial W_\delta(1;x)/\partial x > 0$: $W_\delta(1;x) > 0$ for $x > \tau_a^L$ and $W_\delta(1;x) < 0$ for $x < \tau_a^L$. Therefore, $q = 1$ maximizes $W_\delta(q;x)$ for $x > \tau_a^L$ and $q = 0$ maximizes $W(q;x)$ for $x < \tau_a^L$. Suppose that m is odd, so $a \in I$. Then, because $\tau_m^F \leq \tau_a^L$, the result follows directly from Theorem 7. □

Proof of Theorem 9. It is sufficient to show: (i) Any maximizer of $\widehat{W}_\delta(l)$ satisfies $l \geq \widetilde{l}$; and (ii) for $l \geq \widetilde{l}$, $\widehat{W}_\delta(l)$ is strictly increasing for $l \leq a$ and strictly decreasing for $l \geq a$.

Lemma 6 implies that τ_l^L and τ_l^F are increasing in l, and $\max(I\backslash\{l\})$ is weakly decreasing in l, implying that $\tau^F_{\max(I\backslash\{l\})} \leq \tau_l^L$ for all $l \geq \widetilde{l}$. Therefore, by definition, $\tau_l^* = \tau_l^L$ for $l \geq \widetilde{l}$, implying that τ_l^* is strictly increasing in l for $l \geq \widetilde{l}$. The proof of Theorem 8 shows that $W_\delta(1;x) > 0$ for $x > \tau_a^L$ and $W_\delta(1;x) < 0$ for $x < \tau_a^L$. That establishes (ii). The statement and proof of Lemma 6 are unchanged if the domain of l is expanded to $I \cup \{a\}$. Then Lemma 6 implies $\tau_a^L < \tau_m^L < \overline{x}$, which with $\tau_m^* = \tau_m^L$ implies $\tau_a^L < \tau_m^* < \overline{x}$. Because $W_\delta(1;x) > 0$ for $x > \tau_a^L$, it follows that $\widehat{W}_\delta(m) > 0$. Yet for $l < \widetilde{l}$, the definition of $\widetilde{l}$ implies $\tau_l^L < \tau^F_{\max(I\backslash\{l\})}$, which by definition implies $\tau_l^* = \overline{x}$, implying $\widehat{W}_\delta(l) = 0$. That establishes (i). □

Suppose $\delta = 0$. Then Lemma 2′ becomes Lemma 2, and Lemma 0(iii) implies Lemma 3. The proof of Lemma 6 shows that $\tau_l^L = 0$. With that fact, Theorem 7, as proved for $c = \delta = 0$, becomes Theorem 4. Trivial modifications to the proof of Theorem 8, replacing leader a by any leader, establish Theorem 5.

References

[1] Andreoni, James, "Leadership Giving in Charitable Fund-Raising," manuscript, 2004.

[2] Case, John, *Open-Book Management*, New York: HarperCollins, 1995.

[3] Crawford, Vincent P., and Joel Sobel, "Strategic Information Transmission," *Econometrica* 50:1431-1451 (1982).

[4] Dansereau, Fred, and Francis Yammarino, *Leadership: The Multiple-level Approaches*, London: Jaipress, 1998.

[5] Harsanyi, John., "Games with Incomplete Information Played by 'Bayesian' Players, part II: Bayesian Equilibrium Points," *Management Science* 14:320-334 (1968).

[6] Hermalin, Benjamin E., "Toward an Economic Theory of Leadership: Leading by Example," *American Economic Review* 88:1188-1206 (1998).

[7] Levy, Gilat, and Ronny Razin, "It Takes Two: An Explanation for the Democratic Peace," *Journal of the European Economic Association* 2:1-29 (2004).

[8] Milgrom, Paul., and John Roberts, *Economics, Organization, and Management*, Englewood Cliffs, New Jersey: Prentice Hall, 1992.

[9] Ostrovsky, Michael, and Michael Schwarz, "Equilibrium Information Disclosure: Grade Inflation and Unraveling," manuscript, 2004.

[10] Potters, Jan, Martin Sefton, and Lise Vesterlund, "After You - Endogenous Sequencing in Voluntary Contribution Games," manuscript, 2004.

[11] Prendergast, Canice, "A Theory of 'Yes Men'," *American Economic Review*, 83:757-70 (1993).

[12] Romano, Richard, and Huseyin Yildirim, "Why Charities Announce Donations: A Positive Perspective," *Journal of Public Economics* 81:423-47 (2001).

[13] Rotemberg, Julio J., and Garth Saloner, "Leadership Style and Incentives," *Management Science* 39:1299-1318 (1993).

[14] Varian, Hal, "Sequential Provision of Public Goods, *Journal of Public Economics* 53:165-86 (1992).

[15] Vesterlund, Lise, "The Informational Value of Sequential Fund-raising," *Journal of Public Economics* 87:627-657 (2003).