Toward an Economic Theory of Leadership: Leading by Example*

Benjamin E. Hermalin
University of California at Berkeley

January 14, 1997

Abstract

This paper explores leadership within organizations. Leadership is distinct from authority because following a leader is a voluntary rather than coerced activity of the followers. This paper considers how a leader induces rational followers to follow her in situations when the leader has incentives to mislead her followers.

1 Introduction

Leadership, although long a research topic of other social science disciplines,[1] has been neglected by economics. Economic analyses of organizations have, instead, focused on formal or contractual relationships. Indeed, the players in organizations who would commonly be called leaders, such as managers, are typically modeled as *agents* of other players who are *not* commonly seen as leaders (e.g., shareholders). Such analyses, despite the great insight they have provided on authority and control, do not shed any light on leadership. In particular, they miss what we see as the defining feature of leadership: A

*Helpful comments from Susan Athey, Harold Demsetz, Tom Hubbard, Darrell Williams, and participants at the 2nd Berkeley-Stanford IO*fest*, the UCLA Bergman Workshop in Business Organization and Regulation, and the Winter Econometric Society Meetings in New Orleans are gratefully acknowledged. Proposition 3 is the product of a lengthy discussion with Michael Katz. Section 3.3 was motivated by the insightful comments of Ilya Segal. Conversations with Ruth Konoff, co-president of the Vista-MacGregor PTA, provided some examples. Financial support was provided the U.C. Berkeley Committee on Research, the Walter A. Haas School of Business, and the Harold Furst Chair in Management Philosophy and Values.

[1] Consider, e.g., McGregor (1966).

leader is someone with followers. Following is inherently a *voluntary* activity. Hence, a central question in understanding leadership is how does a leader induce others to follow her.[2]

Even when the leader has *authority*—the power to coerce (directly or indirectly)—such authority is rarely absolute. Certainly the organizational behavior literature still sees a need for managers to encourage and motivate a following (think, e.g., about business school deans and department chairs who must cajole their faculties to take certain actions).[3] Moreover, those people in an organization with authority are not always or solely the leaders. Consider, for instance, that in many academic departments the true leaders are often *not* the department chairs. Consequently, the issues dealt with here apply to a wide range of organizations.

There are, admittedly, many facets to leadership and we can't address all of them here. As economists, we focus on a rational model of leadership; that is, we presume that followers follow because it's in their interest to follow. What could make it in their interest to follow? One answer is that they believe the leader has superior knowledge about what they should do than they have. Leadership is thus, in part, about transmitting knowledge to followers. But this cannot be all there is to leadership: A leader must also convince followers that she is transmitting the correct knowledge; that is, she must convince them that she is not *misleading* them:[4] Consider, for example, the head of some volunteer group like a Parent-Teachers' Association (PTA). Given cutbacks in school funding, there are many activities for volunteers to do. The head of the PTA has an incentive to get volunteers for all of them because that improves her own children's education. Hence, she has an incentive to tell the volunteers that all activities deserve their fullest efforts. Rationally, the volunteers realize that she has these incentives and are, thus, predisposed to *dis*regard her calls to action. The PTA head must,

[2] Handy (1993, p. 97) describes answering this question as one of the "Holy Grails" of organization theory.

[3] See Chapter 4 of Handy (1993) for a survey of current thinking on leadership in organizational behavior.

[4] For example, Kotter (1996), an expert on business leadership, writes:

> Major change is usually impossible unless most employees are willing to help, often to the point of making short-term sacrifices. But people will not make sacrifices, even if they are unhappy with the status quo, unless they think the potential benefits of change are attractive and unless they really believe that a transformation is possible. *Without credible communication*, and a lot of it, employees' hearts and minds are never captured.—Kotter (1996), p. 9 (emphasis added).

therefore, devise a way to convince the volunteers to come out for those activities that are truly the most important. There are two ways that we might see her do that. One is *leader sacrifice*: The leader offers gifts to the volunteers (e.g., a few dozen donuts at a meeting). The volunteers respond *not* because they want the gifts themselves—indeed, the gifts could be public goods that they can enjoy regardless of whether they respond—but because the leader's sacrifice convinces them that she must truly consider this to be a worthwhile activity. The other way to convince the volunteers is *leading by example*: the leader works first, publicly on the activity, thereby convincing the volunteers that she indeed considers this to be a worthwhile activity.

To formalize this intuition and to explore its implications further, we model the leader and her followers as members of a team. The teams model, as formulated by Holmstrom (1982), is well suited to studying leadership. First, because the leader shares in the team's output, she has the necessary incentive to exaggerate the value of effort devoted to the common activity. Second, because the information structure limits the leader's ability to coerce followers, she must induce their voluntary compliance with her wishes.

In the next section, we review the teams problem under symmetric information about the return to effort allocated to a common endeavor. Each team member—*worker*—decides how much effort to invest in the common endeavor versus in private endeavors (including, possibly, his leisure). Under mild conditions, the optimal solution is for the team to share the product of their common endeavor equally. However, as Holmstrom (1982) showed, this solution is only second-best optimal: Because each worker gets only a fraction of the overall (social) return to his effort, he expends less than the first-best level of effort on the common endeavor. That is, because each worker fails to internalize the positive externality his effort has for the team, he expends too little effort. This *teams problem* is, thus, simply an example of the free-riding problem endemic to the allocation of public goods.

In Section 3, we assume that only the leader has information about the return to effort allocated to the common endeavor. Given asymmetric information, the question is how can the leader credibly communicate this information to the rest of the team. We consider two possible solutions. In Section 3.1, we view the problem as a mechanism-design problem. We show that a mechanism that makes *side-payments* among team members a function of the leader's announcement about her information can duplicate the symmetric-information second-best outcome. The leader's side-payment to other workers is increasing in her estimate of the return to effort—she "sacrifices" more the better the state is.

In Section 3.2, we allow the leader to "lead by example"; that is, she expends effort before the other workers. Based on the leader's effort, the other workers form beliefs about the leader's information. We first show, under fairly general assumptions, that leading by example yields an outcome that is *superior* to the symmetric-information outcome. The reason for this surprising conclusion is that the hidden-information problem "counter-acts" the teams problem (free-riding): The need to convince the other workers increases the leader's incentives, so she works harder. In fact, we can reduce her share, but still leave her stronger incentives than under symmetric information, and, thereby, increase the shares of the other workers so that they too work harder. We then proceed to derive what the optimal contract (shares) should be when the leader leads by example. We find that in a small team, she has the smallest share, but in a large team, she has the largest share. Under certain conditions, leading by example dominates symmetric information even when we restrict attention to equal shares.

Because leading by example changes the sequence of moves in the game, it is worth exploring the degree to which its apparent superiority is due to the increased incentives that leading by example represents versus the degree to which it is due to a change in the sequence of moves. We do this in Section 3.3. Under a reasonable class of contracts, the answer is that leading by example's superiority is due solely to the increased incentives. If, however, the class of contracts is expanded to allow "buy-back" contracts—the leader owns 100% of the return to the common endeavor, which she can sell to the other workers after she works—then symmetric information is superior to leading by example. In fact, we show that if the team can employ a series of buy-back contracts it can achieve the first-best outcome under symmetric information (Proposition 8). On the other hand, as we discuss, there are number of reasons to be suspicious of buy-back contracts as a solution to the teams problem. Moreover, buy-back contracts do not work particularly well under *asymmetric* information.

We then consider extensions of this model. As noted above, there is a similarity between the teams problem and a problem of privately providing a public good. We exploit that similarity in Section 4.1 to show why it might be optimal to have public goods provided sequentially rather than simultaneously; thereby explaining phenomena like a charity publicizing the amounts donated by earlier donors to later potential donors. We then turn to the question of multiple leaders, how they compete to induce the team to follow them, and how the team decides which leader to follow. Because there is competition between potential leaders, each leader is induced to be more truthful in her announcements. Consequently, there is less need to sacrifice and, so, competition reduces leader sacrifice. Competition greatly complicates the analysis of leading by example, but we present an example in which leading by example is still a superior system, at least for large teams. We discuss, but do not analyze, two further extensions in Section 4.3. We conclude in Section 5.

As suggested earlier, we are unaware of any earlier economic literature exploring leadership. There is, consequently, little previous literature to which to relate this paper. The idea that one set of players will base their actions on a first-mover because they believe she has information bears some relation to the "herd behavior" or "informational cascades" literature (see, e.g., Banerjee, 1992; Bikhchandani *et al.*, 1992; Scharfstein and Stein, 1990). Unlike that literature, there is no issue here of the followers ignoring their private signals—they have none. More importantly, unlike that literature, the first-mover here has an incentive to induce a following. Finally, unlike that literature, the players here can sign contracts among themselves to affect the transmission of information. Contracting also distinguishes this paper from the Stackelberg-signaling literature in industrial organization (see, e.g., Gal-Or, 1987; Mailath, 1993). Moreover, whereas the leader wishes to convince followers that the state is bad in that literature, here the leader wants to convince them that the state is good. This paper is also related to the growing literature on information transmission within the firm (see, e.g., Prendergast, 1993; Levitt and Snyder, 1996). That literature has tended, however, to view the firm in terms of formal authority and incentive contracts, whereas we are looking at less formal leadership and more voluntary cooperation. A further difference is that literature is about "a fundamental trade-off between inducing workers to tell the truth and inducing them to exert effort" (Prendergast, 1993, p. 769). Our point, in contrast, might be described as the need to convincingly tell the truth can lead to more effort being exerted.

2 Preliminaries

Consider a team with N identical workers indexed by n. Each worker supplies effort e_n. The value to the team of its members' efforts is $V = \theta \sum_{n=1}^{N} e_n$, where $\theta \in [0,1]$ is a stochastic productivity factor. Assume θ is realized after efforts have been supplied. Each worker has the utility function $w - d(e)$, where w is his wage (portion of V) and $d(\cdot)$ is an increasing, convex, and thrice-differentiable function. Assume $d(0) = 0$ and

$$\left[d''(e)\right]^2 + d'''(e) > 0 \text{ for all } e > 0. \tag{1}$$

This last assumption ensures that production will be done by the full team rather than a subset.

The disutility-of-effort function, $d(\cdot)$, can be interpreted as representing the foregone utility of leisure. Alternatively, it can be seen as the worker's foregone profit from reducing his efforts spent on his *private* projects. The latter interpretation is relevant if the role of leadership is seen, in part, to facilitate coordination on *common* projects or objectives.

Assume, keeping with Holmstrom (1982), that although contracts can be written contingent on V, they cannot be written contingent on the team members' efforts. Assume, too, that contracts cannot be contingent on θ *directly* (although they can be contingent on *ex ante announcements* about θ). A contract is a set of contingent wages $\left\{w_n\left(V,\hat{\theta}\right)\right\}_{n=1}^{N}$, where $w_n\left(V,\hat{\theta}\right)$ is worker n's wage when total value is V and the *announced* value of θ is $\hat{\theta}$.[5] Assume that the workers cannot commit to subgame-*imperfect* contracts that would have them forego some of the value (i.e., they cannot use contracts such that $\sum_{n=1}^{N} w_n\left(V,\hat{\theta}\right) < V$). There is no external source of funds, so $\sum_{n=1}^{N} w_n\left(V,\hat{\theta}\right) \not> V$. Hence, attention can be limited to contracts in which $\sum_{n=1}^{N} w_n\left(V,\hat{\theta}\right) = V$ for all V and $\hat{\theta}$. We assume $w_n\left(V,\hat{\theta}\right)$ is differentiable in V for all n and $\hat{\theta}$. Although making assumptions about endogenous functions is not ideal, this assumption is necessary for tractability. Moreover, this is consistent with most real-world compensation plans, in so far as they are continuous and smooth.

Assuming that $w_n\left(\cdot,\hat{\theta}\right)$ is differentiable rules *out* so-called "boss contracts": One worker, the Nth, is the "boss" and her compensation is

$$w_N\left(V;\hat{\theta}\right) = \begin{cases} V \text{ if } V < V\left(\hat{\theta}\right) \\ 0 \text{ if } V \geq V\left(\hat{\theta}\right) \end{cases} .$$

Each of the remaining workers gets

$$w_n\left(V;\hat{\theta}\right) = \begin{cases} 0 \text{ if } V < V\left(\hat{\theta}\right) \\ \frac{V}{N-1} \text{ if } V \geq V\left(\hat{\theta}\right) \end{cases} .$$

That is, the boss gets all the value if value falls below some cutoff; otherwise the value is divided among the other workers. Boss contracts are a partial

[5] Since the workers are risk neutral, random contracts (see Rasmusen, 1987) have no value.

solution to the free-rider problem endemic to team production (Holmstrom, 1982): The threat of losing everything provides extreme incentives to the productive workers ($n = 1, \ldots, N-1$), which prevents their free-riding. The cost is that the boss supplies no effort, so her contribution is lost.

Boss contracts strike us as unrealistic, as we are unaware of any real-world use. As is well known, they are generally not robust to noise (e.g., if $V = \theta \sum_{n=1}^{N} e_n + \nu$, where ν is another stochastic variable). Moreover, and more importantly, they are not coalition-proof: Given that the boss gets nothing if the workers exceed the cutoff, the boss has an incentive to enter into a side contract with one worker to sabotage the team's effort. For instance, the worker can agree to reduce his effort by η, so that $V = V\left(\hat{\theta}\right) - \hat{\theta}\eta$, in exchange for splitting V with the boss. Assuming $N \geq 3$, this is a better deal for the sabotaging worker if η is small. For these reasons, we are comfortable ruling out boss contracts as a solution to the teams problem.

The following lemma will prove useful in what follows:

Lemma 1 *Assume that each worker holds the same beliefs about θ's value conditional on hearing $\hat{\theta}$. For any differentiable contract $\left\{w_n\left(V, \hat{\theta}\right)\right\}_{n=1}^{N}$, there exists an* affine shares contract $\left\{s_n\left(\hat{\theta}\right), t_n\left(\hat{\theta}\right)\right\}_{n=1}^{N}$ *with the following properties:*

- *worker n is paid $s_n\left(\hat{\theta}\right) \cdot V + t_n\left(\hat{\theta}\right)$ when V is realized;*
- $\sum_{n=1}^{N} s_n\left(\hat{\theta}\right) = 1$*;*
- $\sum_{n=1}^{N} t_n\left(\hat{\theta}\right) = 0$*;*
- *this contract induces the same effort from each worker in equilibrium as would $\left\{w_n\left(V, \hat{\theta}\right)\right\}_{n=1}^{N}$; and*
- *this contract yields each worker the same expected utility in equilibrium as would $\left\{w_n\left(V, \hat{\theta}\right)\right\}_{n=1}^{N}$.*

Proof: Let $\tilde{\theta} = \mathbb{E}_\theta\left\{\theta|\hat{\theta}\right\}$. Let $\tilde{e}_n$ be worker n's equilibrium level of effort given $\left\{w_n\left(V, \hat{\theta}\right)\right\}_{n=1}^{N}$. Then it must be that

$$\mathbb{E}_\theta\left\{\frac{\partial w_n}{\partial V}\theta|\hat{\theta}\right\} - d'\left(\tilde{e}_n\right) = 0. \tag{2}$$

Define

$$s_n\left(\hat{\theta}\right) = \frac{\mathbb{E}_\theta\left\{\frac{\partial w_n}{\partial V}\theta|\hat{\theta}\right\}}{\tilde{\theta}}.$$

Since $\sum_{n=1}^N w_n = V$ for all V, it follows that $\sum_{n=1}^N \partial w_n/\partial V = 1$ for all V. Using this fact yields

$$\begin{aligned}\sum_{n=1}^N s_n\left(\hat{\theta}\right) &= \frac{1}{\tilde{\theta}}\sum_{n=1}^N \mathbb{E}_\theta\left\{\frac{\partial w_n}{\partial V}\theta|\hat{\theta}\right\} \\ &= \frac{1}{\tilde{\theta}}\mathbb{E}_\theta\left\{\sum_{n=1}^N \frac{\partial w_n}{\partial V}\theta|\hat{\theta}\right\} \\ &= \frac{\mathbb{E}_\theta\left\{\theta|\hat{\theta}\right\}}{\tilde{\theta}} = 1.\end{aligned}$$

Define

$$t_n\left(\hat{\theta}\right) = \mathbb{E}_\theta\left\{w_n\left(V,\hat{\theta}\right)|\hat{\theta}\right\} - s_n\left(\hat{\theta}\right)\cdot\tilde{\theta}\cdot\sum_{j=1}^N \tilde{e}_j.$$

Since $\sum_{n=1}^N w_n = V$ and $\sum_{n=1}^N s_n = 1$, it follows that $\sum_{n=1}^N t_n = 0$. Under this affine shares contract, each worker chooses e_n to maximize

$$t_n\left(\hat{\theta}\right) + s_n\left(\hat{\theta}\right)\tilde{\theta}\sum_{j=1}^N e_j - d\left(e_n\right).$$

Regardless of his beliefs about his fellow worker's efforts, it is a dominant strategy for him to choose the e_n that solves

$$s_n\left(\hat{\theta}\right)\tilde{\theta} - d'\left(e_n\right) = 0.$$

Since, however, $s_n\left(\hat{\theta}\right) = \mathbb{E}_\theta\left\{\frac{\partial w_n}{\partial V}\theta|\hat{\theta}\right\}/\tilde{\theta}$, this first-order condition must have the same solution as (2). So $\{\tilde{e}_n\}_{n=1}^N$ remains an equilibrium (in fact, it is unique) under this affine shares contract. Finally, since it induces the same effort, it must yield the same expected utility by construction. ∎

In light of Lemma 1, there is no loss of generality in further restricting attention to affine shares contracts when workers hold common beliefs conditional on an announcement $\hat{\theta}$.

Under an affine shares contract each worker's expected utility is

$$\mathcal{U}_n \equiv \tilde{\theta} s_n\left(\hat{\theta}\right)\left(\sum_{m \neq n} e_m + e_n\right) + t_n\left(\hat{\theta}\right) - d\left(e_n\right), \tag{3}$$

where $\tilde{\theta} = \mathbb{E}_\theta\left\{\theta|\hat{\theta}\right\}$. Since increasing s_n raises the marginal benefit of effort, we have

Lemma 2 *A worker's effort is increasing in his share (i.e., $\partial e_n/\partial s_n > 0$).*

We can now solve for the optimal contract when the workers are symmetrically informed.

Proposition 1 *Assume the workers have symmetric information, $\hat{\theta}$ (including, possibly, no information), when they choose their efforts. Then an optimal (second-best efficient)*[6] *contract is an affine shares contract that has $s_n\left(\hat{\theta}\right) = 1/N$ for all n.*

The proof, which is fairly mechanical, can be found in the appendix. Intuitively, if the team were to employ M workers, then the increasing marginal disutility of effort (i.e., $d''(\cdot) > 0$) implies, by Jensen's inequality, that it is optimal to divide the effort evenly among them. From Lemma 2, that means providing them equal shares. The only "difficult" part is to show that $M = N$; that is, all workers are employed. This follows from condition (1).

Given Proposition 1, each worker chooses his e_n to maximize

$$\mathcal{U}_n \equiv \frac{\tilde{\theta}}{N}\left(e + \sum_{j \neq n} e_j\right) + t\left(\hat{\theta}\right) - d'(e).$$

Since

$$\frac{\partial^2 \mathcal{U}_n}{\partial e_n \partial \tilde{\theta}} = \frac{1}{N} > 0,$$

a worker's effort is an increasing function of $\tilde{\theta}$: An increase in $\tilde{\theta}$ raises the marginal benefit of effort, without affecting cost, so more is supplied. To summarize:

Lemma 3 *A worker's effort is increasing in his expectation of θ under an equal shares contract.*

[6] There is still the teams problem (free riding), so the first-best outcome is unattainable.

3 A Hidden-Information Model

Now suppose that one of the workers, *the leader*, gets a private signal concerning θ prior to the expenditure of effort, but *after* contracts have been fixed. Assume, for convenience, that the leader's signal is perfect; that is, she learns what θ will be (this is without loss of generality since the workers are risk neutral).

Suppose, initially, that $s_n = 1/N$ and $t_n = 0$. Suppose that the leader *truthfully* announces her information. Let $e\left(\hat{\theta}\right)$ maximize a worker's utility conditional on believing the announcement $\hat{\theta}$; that is, let it be the solution to

$$\max_e \frac{\hat{\theta}}{N} e - d\left(e\right). \tag{4}$$

From Lemma 3, $e\left(\hat{\theta}\right)$ is increasing in $\hat{\theta}$.[7]

The leader's utility is

$$\frac{\theta}{N}\left(\left(N-1\right) e\left(\hat{\theta}\right) + e\right) - d\left(e\right), \tag{5}$$

which means that she has an incentive to lie to the workers and always announce the maximum possible value for $\hat{\theta}$. For this reason, the workers will rationally disregard her announcement. Valuable information is not utilized, which is sub-optimal relative to a situation in which the leader is induced to announce her information truthfully.

We consider two alternative frameworks for avoiding this inefficient outcome. First, we consider a pure mechanism solution: The contract is contingent on the leader's announcement, $\hat{\theta}$. Second, we consider "leading by example": The leader is allowed to commit to or to choose her effort publicly before the other workers choose their effort (although contracts still cannot be contingent on the leader's effort).

3.1 A Mechanism Design Solution

From the revelation principle, we can restrict attention to mechanisms that induce the leader to tell the truth in equilibrium. Consequently, at the point when the workers choose their effort, they are all symmetrically informed in equilibrium. Proposition 1 therefore applies: If an affine shares contract with equal shares can be found that induces truth-telling, then that contract will be optimal. We now derive such a contract.

[7] Given the assumptions on $d\left(\cdot\right)$, $e\left(\cdot\right)$ is twice differentiable.

Without loss of generality, assign the leader index N (so the remaining workers are $1, \ldots, N-1$). Consider the affine shares contract in which

$$t_N\left(\hat{\theta}\right) = T - \int_0^{\hat{\theta}} z \frac{N-1}{N} e'(z)\, dz; \text{ and} \tag{6}$$

$$t_n\left(\hat{\theta}\right) = \frac{-t_N\left(\hat{\theta}\right)}{N-1} \text{ for } n \leq N-1, \tag{7}$$

where T is an arbitrary constant.

We can now show

Proposition 2 *An affine shares contract in which* $s_n\left(\hat{\theta}\right) = 1/N$ *for all* n *and* $t_n\left(\hat{\theta}\right)$ *is defined by (6) and (7) is optimal (second-best efficient).*

Proof: Given Proposition 1, all we need do is show this contract induces truth-telling by the leader. The leader chooses $\hat{\theta}$ to maximize

$$\frac{\theta}{N}\left(e + (N-1)\, e\left(\hat{\theta}\right)\right) + t_N\left(\hat{\theta}\right) - d(e).$$

The first-order condition is

$$\theta \frac{N-1}{N} e'\left(\hat{\theta}\right) - \hat{\theta} \frac{N-1}{N} e'\left(\hat{\theta}\right) = 0. \tag{8}$$

Clearly, truth-telling—i.e., $\hat{\theta} = \theta$—is a solution. Moreover, since the left-hand side of (8) is positive for $\hat{\theta} < \theta$ and negative for $\hat{\theta} > \theta$, the first-order condition is sufficient as well as necessary. ■

As the proof of Proposition 2 makes clear, $t_N\left(\hat{\theta}\right)$ is decreasing in $\hat{\theta}$. That is, the leader is increasingly "taxed" the better the state is. We call this property *leader sacrifice*. Leader sacrifice corresponds to real-world phenomena in which the leader promises a big party or more vacation time at the end of a big project. Alternatively, she offers inducements to greater effort (participation) by providing donuts and coffee to volunteers who show up for a school's "spruce-up" Saturday. Another example could be a business-school dean who budgets more funds to research computing and faculty facilities than she might prefer in order to convince faculty of the benefits of greater effort in developing an executive education program. In short, we have leader sacrifice whenever a leader promises a *group* reward to convince her team that effort pays big benefits.

Although Proposition 2 requires leader sacrifice, it does not imply that the leader gets less than the other workers. If T, the portion of the side payment not contingent on the announcement, is large enough, then the leader gets more than the other workers.

3.2 Leading by Example

Suppose, now, that the leader can expend effort before the other workers or she can credibly commit to a level of effort.[8] Assume the other workers can observe this, but that contracts cannot be written contingent on the leader's effort. The other workers can, however, make inferences about θ based on the leader's effort.

To begin, consider an affine shares contract satisfying

$$s_N = \frac{1}{1 + N(N-1)} \text{ and } s_n = \frac{N}{1 + N(N-1)} \text{ for } n < N \tag{C1}$$

(recall the leader has index N). Note that the shares are *not* contingent on any announcement by the leader. Assume, too, that neither are the side payments (i.e., $\{t_n\}$). We then have

Proposition 3 (Leading by example I) *Assume the set of possible states is* $[0,1]$*. Then under contract (C1) there exists a separating perfect Bayesian equilibrium in which the workers mimic the leader's effort. Moreover, aggregate welfare (i.e.,* $V - \sum_{n=1}^{N} d(e_n)$*) is greater in this equilibrium than in the pure mechanism-design equilibrium of Proposition 2.*[9]

Proof: We begin with existence. Using the leader's equilibrium strategy, the other workers draw inference $\hat{\theta}$ based on the leader's effort. Their individual best responses are to maximize

$$\hat{\theta}\frac{N}{1 + N(N-1)}e - d(e), \tag{9}$$

[8] As technical matter, there is the question of whether the leader can expend effort both before the others and contemporaneously with them. Whereas this is an important question in the Stackelberg-signaling literature (see, e.g., Mailath, 1993)—in particular, does zero "effort" before guarantee no "effort" with—this is not an issue here because the leader wants to convince the followers that the state is good. Hence, as will become clear, her marginal return to her effort is greater when she expends effort before rather than with the followers. That is, she will expend all her effort before rather than with.

[9] Except when $\theta = 0$. For the sake of brevity, we will not repeat this caveat later.

Let $e_L(\theta)$ denote the leader's equilibrium strategy under (C1). If $e_L(\cdot)$ is invertible, then the other workers can infer θ from the leader's effort (i.e., $\hat{\theta} = \theta$). Let $e^*(\theta)$ maximize (9); i.e.,

$$\theta \frac{N}{1 + N(N-1)} - d'[e^*(\theta)] = 0. \tag{10}$$

Clearly, $e^*(\cdot)$ is monotonic and, hence, invertible. Suppose $e_L(\theta) = e^*(\theta)$. Then it follows from (10) that the other workers' best response is to mimic the leader's effort; i.e., to choose $e = e^*(\theta)$. It remains to check whether $e^*(\theta)$ is the leader's best response to a mimic strategy. Given a mimic strategy, the leader maximizes

$$\theta \frac{1}{1 + N(N-1)} (e + (N-1)e) - d(e). \tag{11}$$

This is identical to (9); hence, conditional on choosing an e in the range of $e^*(\cdot)$, choosing $e^*(\theta)$ is best. What about out-of-equilibrium effort not in the range of $e^*(\cdot)$? Since $e^*(0) = 0$ and $e \geq 0$, the only possible out-of-equilibrium e is $e > e^*(1)$. Since $s_N < s_n$, (10) implies that this could never be optimal for the leader regardless of what beliefs about θ it engendered. So $e^*(\theta)$ is indeed the leader's best response. This establishes that $e_L(\theta) = e^*(\theta)$ and the other workers mimicking the leader is a perfect Bayesian equilibrium.

Now contrast it with the equilibrium of Proposition 2. Aggregate welfare is

$$\sum_{n=1}^{N} (\theta e_n - d(e_n)), \tag{12}$$

which is strictly concave. Let $e^{FB}(\theta)$ maximize (12). As Holmstrom (1982) showed, $e^{FB}(\theta)$ would be the solution if *each* worker were given a 100% share.[10] Since equilibrium effort is an increasing function of a worker's share (recall Lemma 2), it is sufficient to show that

$$\frac{1}{N} < \frac{N}{1 + N(N-1)} < 1, \tag{13}$$

in order to conclude that $e(\theta) < e^*(\theta) < e^{FB}(\theta)$ and, thus, that leading by example yields greater aggregate welfare than the mechanism from Proposition 2. Simple algebra confirms that (13) holds for all $N > 1$. ■

[10] Of course, such a contract is infeasible since the *sum* of shares must equal 100%

The need to signal the state creates an additional incentive for the leader beyond that created by her share of the value created. This can be seen in (11), the expression for her utility: her effort also affects, positively, the efforts of the other workers, which increases her incentives. In fact, her incentives are sufficiently increased that the team can reduce her share, thereby increasing the shares (incentives) for the other workers. Consequently, each member of the team works harder than in the pure mechanism environment of Proposition 2. Because the free-riding endemic to teams means too little effort to begin with, inducing harder work is welfare improving. In fact, since Proposition 2 shows that a pure-mechanism does as well as full information, we have

Corollary 1 *Assume the set of possible states is* $[0,1]$*. Then under contract (C1) aggregate welfare (i.e.,* $V - \sum_{n=1}^{N} d(e_n)$*) is greater with leading by example than under any affine-shares contract given full information.*

Proposition 3 and its corollary constitute an example of two organizational problems—free-riding and impacted information—combining to "cancel each other out." That is, the team does better when it can make the transmission of information costly. This is in the spirit of Caillaud and Hermalin (1993), where introducing a hidden-information agency problem into a signaling problem can be welfare improving because the agency costs raise the *potential* costs of signaling, thereby reducing the production distortion due to signaling.

Admittedly, in addition to adding a signal motive, Proposition 3 also changes the sequence of moves in the game *vis-à-vis* the games considered in Propositions 1 and 2. The importance of this second change to the results will be considered in greater detail in the next sub-section. It should, however, be obvious that changing the sequence of moves has no impact on the earlier results if attention is restricted to affine-shares contracts.

The equilibrium in Proposition 3 depends critically on the assumption that the lower bound of the set of possible states is 0. If the lower bound were greater than zero, then the leader would deviate by choosing an effort below $e^*(\theta)$ for states near the lower bound. Furthermore, Proposition 3 does *not* establish that contract (C1) is optimal: It only establishes that when the leader leads by example, the team is better off than when it is limited to pure-announcement mechanisms as in Proposition 2. These limitations of Proposition 3 suggest a more thorough study of leading by example.

To this end, consider the closed sub-interval $\Theta \subseteq [0,1]$ and let $\underline{\theta} = \min \Theta$. Let s_L be the leader's share and s_W be a worker's share (from

Proposition 1, we know it's optimal to treat the workers equally). Let $\tilde{e}(\theta)$ be the leader's equilibrium strategy (contingent on the contract in place). Since we're interested in the transmission of the leader's information, we will consider separating equilibria only; that is, equilibria in which $\tilde{e}(\cdot)$ is invertible. Conditional on the leader's strategy and effort, e_L, each worker chooses e to maximize

$$s_W \tilde{e}^{-1}(e_L)\, e - d(e). \tag{14}$$

Let $\hat{e}\left[s_W \tilde{e}^{-1}(e_L)\right]$ be the solution. From (14), it is clear that $\hat{e}(\cdot)$ is increasing and differentiable. Anticipating the reaction of the workers, the leader chooses e_L to maximize

$$\theta s_L \left(e_L + (N-1)\,\hat{e}\left[s_W \tilde{e}^{-1}(e_L)\right]\right) - d(e_L).$$

The first-order condition is

$$\theta s_L \left(1 + (N-1)\, s_W \hat{e}'\left[s_W \tilde{e}^{-1}(e_L)\right] \frac{1}{\tilde{e}'\left[\tilde{e}^{-1}(e_L)\right]}\right) - d'(e_L) = 0.$$

In equilibrium, the e_L that solves this first-order condition must equal $\tilde{e}(\theta)$. Making that substitution, we have

$$\theta s_L \left(1 + (N-1)\, s_W \frac{\hat{e}'(s_W \theta)}{\tilde{e}'(\theta)}\right) - d'\left[\tilde{e}(\theta)\right] = 0. \tag{15}$$

The leader's strategy, $\tilde{e}(\cdot)$, is then the solution to this differential equation. As is well-known for signaling games (see, e.g., Mailath, 1987), the initial condition for this differential equation is fixed by the requirement that the "worst" type, $\underline{\theta}$, get her maximum utility conditional on being identified as the worst type; that is, such that $\tilde{e}(\underline{\theta})$ solve

$$\max_e \underline{\theta} s_L \left(e + (N-1)\,\hat{e}(\underline{\theta})\right) - d(e).$$

Consequently, $\tilde{e}(\underline{\theta})$ solves

$$\underline{\theta} s_L - d'(e) = 0. \tag{16}$$

In general, solving the differential equation (15) subject to initial condition (16) is wicked hard. To simplify matters, we will henceforth assume that $d(e) = \frac{1}{2}e^2$. A further limit is we must return to the assumption that $\underline{\theta} = 0$.[11] Under these assumptions we can establish:

Lemma 4 *Assume $d(e) = \frac{1}{2}e^2$ and $\underline{\theta} = 0$. Let the leader's share be s_L. Then her equilibrium strategy is $\tilde{e}(\theta) = k(s_L)\theta$, where*

$$k(s_L) = \frac{s_L + \sqrt{4s_L - 3s_L^2}}{2}. \tag{17}$$

Proof: See the appendix.

We can now consider the optimal contract in this context. Aggregate welfare is

$$\begin{aligned} & \theta\tilde{e}(\theta) + \theta(N-1)\hat{e}(s_W\theta) - \frac{1}{2}\left(\tilde{e}(\theta)^2 + (N-1)\hat{e}(s_W\theta)^2\right) \\ = \quad & \theta^2\left[k(s_L) + (N-1)s_W - \frac{1}{2}\left(k(s_L)^2 + (N-1)s_W^2\right)\right]. \end{aligned}$$

Note that maximizing aggregate welfare with respect to the shares is independent of the value of θ. This means we are free to ignore contracts in which the shares would depend on announcements about θ.[12] Since shares sum to one, $(N-1)s_W = 1 - s_L$; so maximizing aggregate welfare is equivalent to maximizing

$$k(s_L) + 1 - s_L - \frac{1}{2}\left(k(s_L)^2 + \frac{(1-s_L)^2}{N-1}\right) \tag{18}$$

with respect to s_L. The solution has the following properties:

[11] If $\underline{\theta} > 0$, then (15) can be solved using numerical techniques (actually, one must solve the differential equation

$$\mu(e_L)s_L\left(1 + (N-1)s_W\hat{e}'[s_W\mu(e_L)]\mu'(e_L)\right) - d'(e_L) = 0,$$

where $\mu(\cdot) = \tilde{e}^{-1}(\cdot)$, and then invert to work around a singularity at $\underline{\theta}$). When $d(e) = \frac{1}{2}e^2$, the solutions are non-linear near $\underline{\theta}$. This non-linearity, however, tends to die out quickly in θ, leaving a solution that very closely approximates the linear solution found when $\underline{\theta} = 0$ (see Lemma 4). This suggests that the changes in the analysis from allowing $\underline{\theta} > 0$ would be relatively minor when $d(e) = \frac{1}{2}e^2$.

[12] If the shares depended on announcements, then we would have a signaling game with two signals: announcements (shares) and effort. This would greatly complicate the analysis, unless there is no separation on the shares dimension. Without separation, the optimal contract would simply maximize some expected expression. We conjecture that, for at least a representative subset of games, there could be no separation if attention is restricted to "reasonable beliefs" (e.g., as in Cho and Kreps, 1987)—in selecting among contracts, the likely outcome would be that the leader would always choose the contract that gave her the largest share and then separate through her choice of effort.

Proposition 4 (Leading by example II) *Assume* $d(e) = \frac{1}{2}e^2$ *and* $\underline{\theta} = 0$. *Under the optimal contract, the leader works at least as hard as any worker and strictly harder if* $N \geq 3$. *The leader's share,* s_L, *is declining in* N, *but is bounded below by a number approximately equal to .128843. Finally, the leader's share is less than any worker's (i.e.,* $s_L < s_W$*) if* $N \leq 6$, *but greater than his if* $N \geq 7$.

Proof: See the appendix.

Proposition 4 establishes that optimal leading by example can mean that the leader works harder than any individual worker. She works harder, in part, because that is necessary for her to signal her information. In a large team ($N \geq 7$) she also works harder because she gets a larger share of the value created.

That the leader gets a larger share than any other worker in a large team might, at first, seem inconsistent with the intuition given above for the earlier leading-by-example result, Proposition 3. There, recall, the signaling incentive made it possible to increase the incentives (shares) of the other workers without excessively diminishing the leader's overall incentives ("sum" of signaling and share incentives). This effect is still present here. As, however, Proposition 4 demonstrates, a second effect exists: The strength of the signaling incentive depends, in part, on the leader's share. If, for instance, her share were zero, then she would have no incentive to signal.[13] Consequently, because of a need to preserve the signaling incentive, there is a lower bound on the leader's share. Since the average share must tend to zero as the size of the team increases, a lower bound on the leader's share means that eventually her share must come to exceed that of any other worker as the size of the team grows.

That the leader's share is ever less than the other workers' may, at first, seem counterfactual.[14] After all, in most corporations, it is typically upper management (the "leaders") who receive compensation contingent on the corporation's performance, while the production workers receive little contingent compensation. But—to the extent they are teams at all—these are large teams; so this is consistent with Proposition 4. In contrast, consider smaller teams like *maîtres d'* and waiters or floorwalkers and department-store salespeople. In these teams, the workers' compensation is more contingent than the leaders: Waiters typically get a larger share of the tips than

[13] She would also have no incentive to mis-represent θ, but from Proposition 2 that is *not* even the most efficient pure-mechanism, so it must be dominated by leading by example.

[14] Of course, using side-payments (i.e., $\{t_n\}$), it could be that the leader's *overall* compensation is greater.

the *maître d'* and the salespeople get a larger share of the commissions than the floorwalker.[15]

For some teams, such as academic committees and PTAs, the shares are typically fixed and roughly equal. This seems particularly true when the value, V, is non-monetary or indivisible (e.g., a public good). In addition, normative demands for equity could require equal shares. It is therefore worth considering how leading by example works when the shares are fixed at $1/N$ for some reason. Since the leader has the same share incentive as the other workers, but also has a signaling incentive, it follows that she works harder:

Proposition 5 *Assume an equal-shares contract. Then the leader works harder in a separating equilibrium than any individual worker (unless $\theta = \underline{\theta}$, in which case the efforts are the same). Moreover, if $d(e) = \frac{1}{2}e^2$ and $\underline{\theta} = 0$, then leading by example yields greater aggregate welfare than the pure-mechanism of Proposition 2 or any affine-shares contract under full information.*

Proof: The first claim was established in the paragraph proceeding the proposition. Since each worker gets $1/N$ and knows θ when choosing effort, his effort is the same as in Proposition 2 or under the optimal affine-shares contract given full information (Proposition 1). Hence, we need only verify that the leader's equilibrium effort lies between $e(\theta)$ (her full-information best response to a share of $1/N$) and $e^{FB}(\theta)$. Since $d(e) = \frac{1}{2}e^2$, we have $e(\theta) = \theta/N$ and $e^{FB}(\theta) = \theta$. Using Lemma 4,

$$k(1/N) = \frac{1 + \sqrt{(4N-3)}}{2N},$$

which is between $1/N$ and 1 for all $N \geq 2$, so $\theta/N < k(1/N)\,\theta < \theta$. ■

3.3 Sequential Effort and Contract Renegotiation

As noted earlier, one difference between leading by example and the pure-mechanism or full-information cases is that in the latter two cases efforts are chosen simultaneously, whereas the leader plays first in leading by example.

[15] Admittedly, it could be argued that tips and commission are paid on the basis of individual performance rather than group performance so these are not examples of teams. However, it is not uncommon for waiters to pool their tips or salespeople to pool their commissions, particularly when team work is expected.

In this sub-section, we explore the consequences of this change in timing further.

First, as we noted earlier, if we restrict attention to affine-shares contracts, then this change in timing is immaterial: Each worker's choice of effort is a dominant strategy for him or her (i.e., maximizes (3)). Consequently, for a change in timing to matter in a full-information setting or under the pure-mechanism of Proposition 2, the set of possible contracts must be expanded beyond just affine shares.

The alternative is to consider mechanisms that allow the workers to contract *indirectly* on the leader's effort. One possibility would be a mechanism in which all the workers (including the leader) would make announcements about the leader's effort to a third party (e.g., a court), who would then make payments based on a previously-agreed-to schedule for mapping announcements into payments. As, however, Hermalin and Katz (1991, 1993) have pointed out, relationships such as these are not really governed by contracts that use the courts as direct revelation mechanisms. They offer a number of persuasive reasons why this is.[16] We, therefore, follow their approach by restricting attention to contracts contingent on verifiable signals. Like them, however, we allow these contracts to be renegotiated to incorporate new information (e.g., the leader's choice of effort).

The only verifiable information is the final value of the team's efforts (i.e., V). The parties can exploit this by using "buy-back" contracts like those put forth by Demski and Sappington (1991): the leader initially owns 100% of the value, but after she has chosen her effort, the rest of the team can "buy" it back from the leader at a preset price (the parties can also employ *upfront* transfers to deal with the distribution of the surplus). However, as Edlin and Hermalin (1996) point out, these buy-back contracts are not renegotiation proof. We will, therefore, consider a renegotiation-proof variant.

Working backwards in time, workers $1, \ldots, N-1$ are symmetrically informed at the time they buy back "the firm" (i.e., the rights to V); hence, Proposition 1 implies that these workers will work under a contract in which each worker's share is $\frac{1}{N-1}$. Since our focus is on separating equilibria, we may as well assume these workers know θ when choosing their efforts. Let

[16] These reasons include that these mechanisms often admit multiple equilibria; they require the parties to commit to destroy surplus off the equilibrium path, which is subgame imperfect (i.e., subject to renegotiation); and they are particularly sensitive to small pertubations in the information structure. In addition, since actual court cases differ greatly from sending simple messages and being rewarded according to some preset payment schedule, modeling the court in this way is too artificial to be insightful about real life.

$\varepsilon(\theta)$ be a worker's effort conditional on these workers' owning "the firm"; that is, $\varepsilon(\theta)$ maximizes

$$\frac{\theta}{N-1}e - d(e).$$

The value of the firm to these workers is, thus,

$$(N-1)(\theta\varepsilon(\theta) - d[\varepsilon(\theta)]) + \theta e_L,$$

where e_L is the leader's effort. If these workers don't buy back the firm from the leader, then they won't expend any effort and, consequently, the value of the firm to the leader will be just θe_L. It follows that the surplus from trade is

$$(N-1)(\theta\varepsilon(\theta) - d[\varepsilon(\theta)]). \tag{19}$$

Rather than a preset buy-back price, which would be subject to renegotiation, we assume that there is no agreed upon price and the parties simply bargain over a price after the leader has expended her effort. There are many bargaining games that we could use, but for concreteness we will assume that the bargaining yields a price of

$$\beta \cdot (N-1)(\theta\varepsilon(\theta) - d[\varepsilon(\theta)]) + \theta e_L,$$

where $\beta \in [0,1)$; that is, the leader gets β of the surplus from trade plus the value of the firm prior to the other workers' efforts. The leader's utility (gross of any upfront transfers) is, thus,

$$\theta e_L + \beta \cdot [(N-1)(\theta\varepsilon(\theta) - d[\varepsilon(\theta)])] - d(e_L).$$

It follows that the leader will choose $e_L = e^{FB}(\theta)$. The other workers' aggregate utility (gross of any upfront transfers) is

$$(1-\beta) \cdot (N-1)(\theta\varepsilon(\theta) - d[\varepsilon(\theta)]) > 0.$$

To summarize, we have

Proposition 6 *Assume full information and that only the leader chooses effort first. Then the optimal (second-best efficient) outcome is the leader supplies effort $e^{FB}(\theta)$ and each worker supplies effort $\varepsilon(\theta)$. Given constant-shares bargaining (i.e., β independent of the leader's effort), this outcome can be obtained by initially giving ownership of the returns to the leader and then allowing the other workers to bargain back for the returns after the leader has committed her effort.*

Comparing the outcome in this proposition with leading by example yields

Proposition 7 *Assume full information and that the leader chooses effort first. Then the resulting equilibrium outcome yields greater aggregate welfare than does contract (C1) in Proposition 3. If* $d(e) = \frac{1}{2}e^2$ *and* $\underline{\theta} = 0$*, then this outcome yields greater aggregate welfare than the optimal contract in Proposition 4.*

Proof: Consider the first claim. The leader's effort is clearly more optimal. Moreover, since

$$1 > \frac{1}{N-1} > \frac{N}{1 + N(N-1)},$$

the workers' efforts are also more optimal.

Consider the second claim. The leader's effort is clearly more optimal ($k(s_L) < 1$). For $N \geq 7$, $s_W < \frac{1}{N} < \frac{1}{N-1}$. For $N \leq 6$, Table 1 in the Appendix establishes that $s_W \leq \frac{1}{N-1}$. ■

Indeed, as should be intuitively clear, if we let all the workers work in sequence, then we can achieve the first-best outcome using a series of buy-back arrangements:

Proposition 8 *Assume full information and that the workers can choose and observe effort sequentially. Then the first-best outcome is attainable using a series of buy-back arrangements assuming all bargaining is constant-shares bargaining.*

Since this last result is tangential to our focus here, we've relegated its proof to the appendix.

Consequently, if we are willing to expand the set of feasible contracts to allow for buy-back arrangements, then the conclusion of the previous sub-section that leading by example dominates full information no longer holds.

Duplicating the results of Propositions 6 and 7 given asymmetric information is made difficult by the fact that were the leader to lie about θ, then the bargaining over ownership of the firm would be bargaining under asymmetric information. Consequently, as is well known, the results will be quite sensitive to the extensive form assumed for the bargaining game. For instance, if we restrict renegotiation to a mechanism based on the leader's announcement, then we cannot duplicate the full-information results: Since the seller (i.e., the leader) has private information about the buyer's (i.e., the other workers') valuation, it follows from Proposition 2 of Hermalin and Katz (1993) that the full-information solution is *un*attainable for the class of mechanisms we are allowing here.

Proposition 9 (Hermalin and Katz) *Under* asymmetric *information, the full-information solution is unattainable using buy-back mechanisms.*

Alternatively, suppose we restrict attention to the following variant of leading by example and buy-back mechanisms: The leader initially owns the firm. After she expends effort, e_L, the other workers can make the leader a take-it-or-leave-it offer for the firm at a price p. Suppose, too, as a strong assumption, that no further negotiations are possible. Let $\tilde{e}(\theta)$ be the leader's strategy. Assuming a separating equilibrium, the workers will bid $e_L \cdot \tilde{e}^{-1}(e_L)$ for the firm, which the leader must accept in a second-best efficient equilibrium. The leader's payoff is, thus,

$$e_L \cdot \tilde{e}^{-1}(e_L) - d(e_L). \tag{20}$$

The first-order condition for her maximization program is

$$\tilde{e}^{-1}(e_L) + \frac{e_L}{\tilde{e}'\left[\tilde{e}^{-1}(e_L)\right]} - d'(e_L) = 0;$$

or, using the fact that $e_L = \tilde{e}(\theta)$ in equilibrium,

$$\theta + \frac{\tilde{e}(\theta)}{\tilde{e}'(\theta)} - d'\left[\tilde{e}(\theta)\right] = 0. \tag{21}$$

As before, the initial condition for this differential equation is

$$\underline{\theta} - d'\left[\tilde{e}(\underline{\theta})\right] = 0. \tag{22}$$

Note that unless $\theta = \underline{\theta}$, (21) implies $\tilde{e}(\theta) > e^{FB}(\theta)$. Equation (21) is, however, only one incentive-compatibility constraint in play here: There is also the constraint that the leader not "steal the firm." If the leader chooses an $e_L < \tilde{e}(\theta)$, she will end up owning the firm (the other workers will bid too low), which yields her

$$\theta e_L - d(e_L). \tag{23}$$

She maximizes this by choosing $e_L = e^{FB}(\theta)$. But $e^{FB}(\theta)$ is less than $\tilde{e}(\theta)$ for all $\theta > \underline{\theta}$, which by definition of $e^{FB}(\theta)$ means that (23) exceeds (20): The leader would steal the firm if the workers believe she is playing strategy $\tilde{e}(\cdot)$. This shows that a separating equilibrium is impossible—the two incentive-compatibility constraints cannot be simultaneously met. To summarize:

Proposition 10 *Assume the leader leads by example and the team is using a non-renegotiable buy-back mechanism that gives all the bargaining power to the workers. Then no separating equilibrium exists.*

In light of this proposition, it is impossible to exploit fully the leader's information in this setting.

It might at first seem that Proposition 10 is a consequence of having to provide incentives for the leader not to "steal the store"; which suggests that the problem could be alleviated by giving the leader a bigger share of the surplus from selling—e.g., give her all the bargaining power. This intuition, however, is incomplete at best: By giving the leader a bigger share of the surplus, her incentives to signal are increased. This, in turn, increases the distortions in her effort *vis-à-vis* the first best (recall the $\tilde{e}(\theta)$ that solves (21) *exceeds* $e^{FB}(\theta)$). These distortions can increase so rapidly in her share of the surplus, that she still prefers to steal the store (i.e., deviate if the workers believe she is playing $\tilde{e}(\theta)$ by playing $e^{FB}(\theta)$ and retaining ownership). For instance, it can be shown for the case in which $\underline{\theta} = 0$ and $d(e) = \frac{1}{2}e^2$ that no separating equilibrium exists even if the leader has all the bargaining power.[17]

Even if it were possible to extend the results of Propositions 6–8 to cover asymmetric information, there are a number of potential objections to buy-back mechanisms in this context. First, in many team situations it is difficult to see how the team could literally "sell the store" among members. A dean, for example, can't be given ownership of a business school, nor can she sell it to the faculty. Admittedly, the store doesn't actually have to be "sold" to operationalize a buy-back mechanism, but it may be hard to establish ownership rights to a stream of returns otherwise. Indeed, for the reasons discussed in connection with Proposition 5, the team may be limited to very simple contracts (e.g., fixed shares). Another objection is that the analysis presented so far ignores bargaining costs. Given that the other workers have to coordinate their bargaining position *vis-à-vis* the leader, these costs could

[17] For the leader to signal credibly, she must play $\tilde{e}(\theta) = \bar{k}\theta$, where

$$\bar{k} = \frac{1}{2}N + \frac{1}{2}\sqrt{N^2 + 8N - 12}.$$

Her utility from playing $\tilde{e}(\theta)$ is

$$\theta^2\left(\bar{k} + 1 - \frac{1}{2(N-1)^2} - \frac{1}{2}\bar{k}^2\right).$$

Her utility from stealing the store is $\frac{1}{2}\theta^2$. Simple algebra confirms that this is greater than her utility from playing $\tilde{e}(\theta)$.

be considerable. Finally, unlike leading by example, buy-back mechanisms are not robust to other sources of asymmetric information. For instance, suppose that

$$V = \theta \sum_{n=1}^{N} e_n + \nu,$$

where ν is a stochastic factor known only to the leader. Then leading by example would still work (as is readily verified), but the previously identified problems with buy-back mechanisms under asymmetric information would resurface.

Although buy-back mechanisms are clearly a powerful means of dealing with the free-riding problem endemic to teams problems, it is unclear, for the reasons enumerated above, that they are particularly well suited to the problem of a leader disclosing valuable information; indeed, as Proposition 10 shows, buy-back mechanisms can exacerbate the problem of information revelation.

4 Extensions

4.1 Private Provision of Public Goods

A teams problem is similar to a public-goods problem: The workers' efforts can be seen as private contributions to a public good. There are, admittedly, differences. In particular, a team can exclude some members from receiving the public good (a necessity for buy-back mechanisms) and it can adjust the members' share (as in contract (C1) and Proposition 4), whereas the public-good problem stems from an inability to exclude people from the public good. On the other hand, Proposition 5 translates immediately into the public-good context: Let a citizen's utility from the public good, V, and the private good, x_n, be

$$U(V) + W(x_n),$$

where

$$U(V) = \frac{1}{N} V, \tag{24}$$

$$W(x_n) = -d\left[\Phi(\bar{x} - x_n)\right], \tag{25}$$

and a citizen transforms $\bar{x} - x_n$ of his endowment, $\bar{x}$, of the private good into a portion of the public good according to the production process $\theta\Phi(\bar{x} - x_n)$. Imagine that there is a "lead citizen," who learns θ and who donates before the other citizens. Then, from Proposition 5, we have:

Corollary 2 *Consider a public-goods problem in which each citizen's utility for the private good,* x_n*, and the public good,*

$$V = \theta \sum_{n=1}^{N} \Phi\left(\bar{x} - x_n\right),$$

is given by expressions (24) and (25). Suppose $\Phi(\cdot)$ *is strictly increasing. Then the "lead citizen" donates more in a separating equilibrium than any individual citizen (unless* $\theta = \underline{\theta}$*, in which case the donations are the same). Moreover, if* $d(e) = \frac{1}{2}e^2$ *and* $\underline{\theta} = 0$*, then leading by example yields greater aggregate welfare than simultaneous donations under full information.*

Many charitable and not-for-profit enterprises (e.g., hospitals, orchestras, business schools) report a list of their big donors and the amounts they donate when soliciting additional donations.[18] This is particularly true of lead or naming donations at the start of a capital campaign. Corollary 2 offers an explanation for this behavior: Knowing that their donations will be publicized, lead donors realize that their donations will influence later donors, who will interpret their donations as signals of the worthiness of the charity or non-for-profit enterprise. This induces lead donors to donate more than they would were donations solicited simultaneously (or previous donations kept secret).

Corollary 2 reaches a different conclusion than Varian (1994), where sequential donations prove to be *worse* than simultaneous donations. In Varian, the lead donor exploits her Stackelberg position by donating *less*; that is, she is able to free-ride more. There are two reasons for this difference. First, in Varian, the utility functions are different. There, $U(\cdot)$ is concave instead of linear. The concavity of $U(\cdot)$ means that the marginal return to a second donor's donation is *decreasing* in a first donor's donation; conversely, by donating less, a first donor raises the marginal return to a second donor's donation, inducing him to donate more. Here, in contrast, $U(\cdot)$ is linear, so a first donor's donation has no impact on the marginal return to a second donor's donation. Consequently, Varian's Stackelberg free-riding does not arise. The second difference between this paper and Varian's is that the lead donor has private information, which gives the first donor incentives not present in Varian.

[18] Or at least a ball-park figure on the amount donated.

4.2 Who's Followed?

So far we've assumed that there is only one leader, namely the worker who is endowed with the information about the stochastic productivity parameter θ. What if, in contrast, more than one worker was endowed with information about how the team should allocate its efforts? How do would-be leaders induce others to follow?

There are many ways in which we could formalize these questions. Here we pursue but one, leaving the others for future research. There are two possible projects, 1 and 2, that the team could pursue. The value from pursuing project i is

$$V_i = \theta_i \sum_{n=1}^{N} e_n, \tag{26}$$

where θ_i is the productivity parameter that applies to project i. Largely for convenience, assume that θ_1 and θ_2 are independently, but identically distributed.

Since (26) is linear in $\{e_n\}$, the team would want, in a first-best world, to devote its efforts to one project or the other, but not divide them between the projects.[19] In a second-best world with free-riding, the team might want to divide its efforts if both V_1 and V_2 are verifiable: By splitting the team between the projects, each worker can be given a larger share of his own effort (e.g., $2/N$ versus $1/N$), which increases his effort (recall Lemma 2). If θ_1 and θ_2 are not too different, this increase in effort offsets the lower productivity of one of the projects. Although splitting the team for this reason raises some interesting issues, they are tangential to our interest in information and leadership.[20] Consequently, we assume that only the *sum* of V_1 and V_2 can be verified. Hence, even in a second-best world, it is optimal to devote effort to only one project.

There are two *potential* leaders, L_1 and L_2. Potential leader L_i learns θ_i perfectly. Her knowledge of θ_i is her private information. When announcing θ_i, a potential leader has two incentives now. As before, she has an incentive to overstate θ_i, since this will induce more effort from the rest of the team. But now, if she overstates θ_i, she risks that the worse project will be pursued (i.e., project i is pursued when $\theta_i < \theta_j$). This second incentive does not, however, offset the first:

[19] Unless $\theta_1 = \theta_2$, but this is a probability zero event.

[20] Moreover, to some extent, a split team can be dealt with using the previous analysis.

Proposition 11 *Assume an equal-shares contract.*[21] *Then potential leader L_i's best response to truth-telling by the other potential leader and credulity by the workers is to overstate θ_i. Consequently, there is no equilibrium in this setting in which both potential leaders tell the truth as a pure strategy.*

Proof: See the appendix.

Intuitively, the risk of the wrong project being chosen by slightly overstating θ_i and the cost if it is are both close to zero. However, exaggerating will induce harder effort by the other workers if project i is chosen. The benefit of exaggerating therefore exceeds its negligible cost.

As in the single-leader case, we can induce truth-telling using a mechanism: Let $F(\cdot)$ be the marginal distribution function for θ_i and let $f(\cdot)$ be the corresponding density function. Set the transfers as follows.

$$
\begin{aligned}
t_{L_1}\left(\hat{\theta}_1, \hat{\theta}_2\right) &= T_1 - \int_0^{\hat{\theta}_1} \frac{z}{N}(N-1)\, e'(z)\, F(z)\, dz, \\
t_{L_2}\left(\hat{\theta}_1, \hat{\theta}_2\right) &= T_2 - \int_0^{\hat{\theta}_2} \frac{z}{N}(N-1)\, e'(z)\, F(z)\, dz, \text{ and} \qquad \text{(C2)} \\
t_W\left(\hat{\theta}_1, \hat{\theta}_2\right) &= -\frac{t_{L_1}\left(\hat{\theta}_1, \hat{\theta}_2\right) + t_{L_2}\left(\hat{\theta}_1, \hat{\theta}_2\right)}{N-2},
\end{aligned}
$$

where $e(\cdot)$ is a worker's or leader's dominant strategy under an equal shares contract given he or she believes the state is θ. Assume equal shares. Then we have

Proposition 12 *An equal-shares contract in which transfers are given by (C2) and project i is pursued if and only if $\theta_i \geq \theta_j$ yields a second-best efficient equilibrium.*

Proof: See the appendix.

As in the single-leader case, we have *leader sacrifice*; except here we have both (potential) leaders sacrificing. From (C2), the better a leader's project is, the more she sacrifices. If $T_1 = T_2$, the leader who sacrifices most—gives the workers the larger "gift"—induces the workers to follow her. That is, we have a "leadership struggle" in which would-be leaders compete through their gifts to potential followers.

Although there is a leadership struggle, each leader gives a smaller gift than she would give in the single-leader case; that is,

[21] Which would be second-best efficient given truthful revelation.

Proposition 13 *Fixing $T = T_1 = T_2$, a leader gives a smaller gift when there is another potential leader than when she is the only leader.*

Proof: Compare (C2) with (6). ■

It might at first seem counter-intuitive that a leadership struggle lessens the amount that a potential leader must give to induce a following. But remember a leader's objective is to convince followers of the state, θ. When she has competition, her incentive to lie is reduced (since effort could be allocated to the wrong project). Consequently, the amount she must give the workers to convince them of the value of θ is less than it would be absent competition.

What about leading by example? To answer this, we need to be clear about what is entailed by a potential leader expending effort before the other workers. There are many possible assumptions we could make, but we will consider just one here: The potential leaders expend effort first, each on her project. They cannot expend more effort later. Effort expended on one project cannot be redirected later to another project. We will focus on symmetric strategies for the leaders. Let $\tilde{e}(\theta_i)$ be L_i's effort conditional on her project's productivity parameter. Given our interest in information revelation, we consider only invertible functions $\tilde{e}(\cdot)$.

The workers will follow the more productive project; that is, project i if

$$\tilde{e}^{-1}(e_{L_i}) \geq \tilde{e}^{-1}\left(e_{L_j}\right).$$

Let $\hat{e}\left[s_W \tilde{e}^{-1}(e_{L_i})\right]$ be the e that maximizes

$$s_W \tilde{e}^{-1}(e_{L_i})\, e - d(e).$$

Potential leader i's expected utility as a function of her effort e_{L_i} is, thus,

$$s_L \theta_i e_{L_i} + s_L \int_0^1 z \tilde{e}(z) f(z)\, dz + s_L \theta_i (N-2) F\left[\tilde{e}^{-1}(e_{L_i})\right] \hat{e}\left[s_W \tilde{e}^{-1}(e_{L_i})\right]$$
$$+ \int_{\tilde{e}^{-1}\left(e_{L_i}\right)}^1 s_L z (N-2)\, \hat{e}(s_W z) f(z)\, dz - d(e_{L_i}).$$

On the equilibrium path, L_i's first-order condition is

$$s_L \theta_i + s_L \theta_i (N-2) F(\theta_i) \frac{s_W \hat{e}'(s_W \theta_i)}{\tilde{e}'(\theta_i)} - d'\left[\tilde{e}(\theta_i)\right] = 0. \tag{27}$$

As before the initial condition is

$$s_L \underline{\theta} - d'\left[\tilde{e}(\underline{\theta})\right] = 0. \tag{28}$$

The differential equation (27) admits no analytic solution in general and we have been unable to find an analytic solution for any reasonable specific parameterization. This makes comparing the welfare properties of leading by example to those of full information or mechanism (C2) difficult. We can, however, derive some properties concerning the solution.

Proposition 14 *A potential leader's equilibrium strategy, $\tilde{e}(\theta)$, satisfies the following properties:*

- $\tilde{e}'(\theta) > 0$*; and*
- *Compared to a worker with an equal share who follows her, a potential leader works harder than that worker.*

Proof: See the appendix.

The second point of Proposition 14 corresponds to the earlier instances of leading by example. Once again the need to signal to the workers provides additional incentives to the leaders beyond that provided by their shares. Hence, leading by example still offsets, to some extent, the free-riding problem endemic to team production.

There is, however, a confounding problem with leading by example in this context: One leader is working on the wrong project. Although her effort is not completely wasted, its marginal return is less than if she were working on the more productive project. Whether this problem outweighs the benefits of leading by example depends on how likely it is that $\theta_i \ll \theta_j$. To get some idea about this, we assumed that θ was distributed uniformly, we fixed all shares at $1/N$, we set $d(e) = e^2/2$ and $\underline{\theta} = 0$, and we approximated $\tilde{e}(\theta)$ by $\hat{k}\theta$, where

$$\hat{k} = \frac{3}{4N} + \frac{1}{4}\sqrt{\frac{4}{N} - \frac{7}{N^2}}.$$

Comparing specific numerical solutions of (27) to $\hat{k}\theta$ indicates that the latter *understates* $\tilde{e}(\theta)$.[22] Hence, our analysis overweights the problem of one potential leader working on the wrong project. Under full-information (no leading by example), expected aggregate welfare is

$$EW^{FI} = \mathbb{E}\left\{\max(\theta_1, \theta_2)^2\right\} \times \frac{2N-1}{2N}$$

[22] The *Mathematica* program used for solving (27) numerically is available upon request. The function $\hat{k}\theta$ is the average of two linear functions that bound $\tilde{e}(\theta)$.

$$
\begin{aligned}
&= \frac{2N-1}{2N}\int_0^1 2z^3 dz \\
&= \frac{2N-1}{4N}.
\end{aligned}
$$

Under leading by example, expected aggregate welfare is

$$
\begin{aligned}
EW^{LBE} &= \mathbb{E}\left\{\max\left(\theta_1,\theta_2\right)^2\right\}\frac{(2N-1)(N-2)}{2N^2} + 2\mathbb{E}\left\{\theta^2\right\}\left(\hat{k}-\frac{1}{2}\hat{k}^2\right) \\
&= \frac{(2N-1)(N-2)}{4N^2} + \frac{2}{3}\left(\hat{k}-\frac{1}{2}\hat{k}^2\right).
\end{aligned}
$$

The difference is

$$
EW^{LBE} - EW^{FI} = -\frac{2N-1}{2N^2} + \frac{2}{3}\left(\frac{3}{4N} + \frac{\sqrt{\frac{4}{N}-\frac{7}{N^2}}}{4} - \frac{1}{2}\left[\frac{3}{4N} + \frac{\sqrt{\frac{4}{N}-\frac{7}{N^2}}}{4}\right]^2\right).
$$

Tedious, but straightforward algebra, reveals that this difference is positive for all $N \geq 5$. This supports the following conjecture:

Conjecture 1 *For a large enough team, the expected gains from leading by example outweigh the expected loss from having one leader work on the wrong project.*

Intuitively, having a leader be just another worker offers little gain when the team is large and her effort is, therefore, a small proportion of all effort. Put another way, the relative expected loss from her working on the wrong project is small. However, the guaranteed benefit of giving her additional incentives (a need to signal) is still relatively great.

4.3 Future Extensions

Consistent with the title of this paper, we believe that this is the beginning, not the end, of a line of research. Here, we briefly discuss potential future extensions of this work (in addition to those identified earlier) and some of the issues involved with them.

4.3.1 Research Effort

So far we've assumed that the leader receives her information about θ for free. A logical extension would be to assume that she must expend costly effort or make some other investment to acquire this information. There are many ways this could be modeled: she acquires a signal only if she puts in "research" effort above some threshold; her probability of acquiring a signal is an increasing function of her research effort; or the precision of her signal is an increasing function of her research effort. One open question is how the contracts of the earlier sections would need to be adjusted to give her incentives to invest research effort (since she never realizes 100% of the gain from the information, it is unlikely that she can be induced to expend the first-best level of research effort).

Another set of issues deals with changes in the information structure. Given our focus on separating equilibria, one important issue is whether the support of θ differs depending on the leader's research effort. A second issue has to do with the disutility of effort$s$ function. If it is additively separable between research effort and productive effort, then there is no problem—provided the support of θ is independent of research effort, the analysis of Sections 3.1 and 3.2 still applies (although contracts might need to be adjusted to provide appropriate incentives for research effort). If it is not additively separable, then the leader's type has two dimensions: her research effort, which determines her marginal disutility of productive effort, and her knowledge of θ. This would, then, greatly complicate the analysis.[23]

4.3.2 Leadership Ability

We have so far assumed that the leader(s) observe θ perfectly. For our purposes, this is without loss of generality given that the workers are risk neutral: Let q be the leader's best estimate of θ conditional on her private information and the prior distribution of θ. Then the preceding analysis still applies with q replacing θ.

This is not to say, however, that there is no cost to an imprecise estimate—effort is increasingly mis-allocated in expectation the less precisely θ is estimated. Suppose, therefore, that different potential leaders have different abilities to estimate θ; e.g., each leader's ability, α, is a measure of the precision of her estimates. In a dynamic model, one issue would be to infer the leader's ability and to replace her if her *estimated* precision falls below some cutoff.

[23] See Fudenberg and Tirole (1990) and Ma (1994) for discussions of signaling and screening issues when an agent's type is effort (e.g., as here, prior research effort).

5 Conclusion

We have seen in this paper that it is possible to construct an economically rigorous model of leadership; one that captures the fundamental feature of leadership, namely that leaders have followers and following is a voluntary activity. To limit the scope for authority, we modeled leadership within a team. The leader had better information about the marginal return, θ, to effort committed to a common endeavor. Her problem was to convey this information accurately given that the other workers in the team recognized that she had an incentive to overstate θ. We explored two ways that she could do this. First, she could *sacrifice*; that is, make a side-payment to the other workers that was *not* a function of their efforts. The more she was willing to sacrifice, the greater the workers believed θ to be. We showed that leader sacrifice allowed the team to do as well as it could under symmetric (full) information. Second, she could *lead by example*; that is, commit effort first to signal information about θ. The harder she worked, the harder the other workers worked. We showed that leading by example could be *superior* to the symmetric-information outcome: The need to signal gave the leader additional incentives, which helped to offset the teams problem (too little effort). The change in timing underlying leading by example, led us to explore sequential play in teams. We showed that if the team could use "buy-back" contracts, then it could do better under symmetric information with sequential play than with simultaneous play. Indeed, it was possible for the team to achieve the first best. Finally, we considered two extensions: exploiting the similarity between the teams problem and the private provision of a public goods to re-examine whether simultaneous or sequential provision of a public good was optimal; and exploring the consequences of having two potential leaders and a leadership struggle. As noted previously, there is much work that remains to be done. We are confident, however, that continued study by economists will help to develop an economic understanding of the less formal aspects of organization, such as leadership.

Appendix: Proofs

Proof of Proposition 1: Let $\tilde{\theta}$ be the expected value of θ conditional on $\hat{\theta}$. From Lemma 1, attention can be limited to affine shares contracts.

Collectively, the workers seek to

$$\max_{\{s_n\}} \sum_{n=1}^{N} \left(\tilde{\theta} e\left(s_n\right) - d\left[e\left(s_n\right)\right]\right) \tag{29}$$

$$\text{subject to } \sum_{n=1}^{N} s_n = 1, \tag{30}$$

where $e(\cdot)$ is defined by the individual worker's first-order condition:

$$\tilde{\theta} s_n - d'(e) = 0. \tag{31}$$

The first-order condition for (29) is

$$\tilde{\theta} e'\left(s_n\right) - d'\left[e\left(s_n\right)\right] \cdot e'\left(s_n\right) - \lambda = 0 \text{ for all } n, \tag{32}$$

where λ is the Lagrange multiplier on (30).

We'll first show that the unique solution to (32) is symmetric. Using (31), (32) can be rewritten as

$$\tilde{\theta}\left(1 - s_n\right) e'\left(s_n\right) - \lambda = 0.$$

This has a unique solution if $(1 - s_n) e'(s_n)$ is monotonic in s_n. Differentiating it yields

$$-e'\left(s_n\right) + \left(1 - s_n\right) e''\left(s_n\right). \tag{33}$$

From (31):

$$e'\left(s_n\right) = \frac{\tilde{\theta}}{d''(e)} \text{ and } e''\left(s_n\right) = -\frac{\tilde{\theta} d'''(e)}{\left[d''(e)\right]^2} \cdot e'\left(s_n\right).$$

Hence, the sign of (33) is the same as the sign of

$$-1 - \left(1 - s_n\right) \frac{\tilde{\theta} d'''(e)}{\left[d''(e)\right]^2}.$$

Since $\tilde{\theta} \in [0,1]$ and $s_n \in [0,1]$, assumption (1) ensures that this last expression is negative; so $(1 - s_n) e'(s_n)$ is monotonic. Consequently, (32) has only one solution; that is, $s_n = s_m = 1/N$ for all n and m.

Finally, we show that the second-order conditions are met. It's sufficient to show that $\tilde{\theta} e(s_n) - d[e(s_n)]$ is concave in s_n: the second derivative is

$$\tilde{\theta} e''\left(s_n\right) - d''\left[e\left(s_n\right)\right] \cdot \left[e'\left(s_n\right)\right]^2 - d'\left[e\left(s_n\right)\right] e''\left(s_n\right).$$

Using (31) this can be rewritten as

$$-\tilde{\theta}\left(1-s_n\right)\frac{\tilde{\theta}d'''\left(e\right)}{\left[d''\left(e\right)\right]^2}e'\left(s_n\right)-\tilde{\theta}e'\left(s_n\right).$$

Using (1) it can be shown that this is negative. ■

Proof of Lemma 4: From (14) and (16), it follows that $\hat{e}(\cdot)$ is the identity function (i.e., $\hat{e}(x)=x$) and $\tilde{e}(\underline{\theta})=0$. Note the latter coincides with the definition given in the statement of the Lemma. Since shares sum to one, $(N-1)\,s_W=1-s_L$. We can therefore rewrite (15) as

$$\theta s_L\left(1+\left(1-s_L\right)\frac{1}{\tilde{e}'\left(\theta\right)}\right)-\tilde{e}\left(\theta\right)=0.$$

We need only confirm that $k\left(s_L\right)\theta$ solves this differential equation:

$$\begin{aligned}\theta s_L\left(1+\frac{1-s_L}{k\left(s_L\right)}\right)-k\left(s_L\right)\theta &= 0 \text{ if}\\ k\left(s_L\right)^2-s_Lk\left(s_L\right)-s_L\left(1-s_L\right) &= 0.\end{aligned}$$

Simple algebra confirms that the $k(\cdot)$ defined in (17) solves this last equation. ■

Proof of Proposition 4: As preliminaries, we establish two claims.
Claim 1: If $k\left(s_L\right)>s_W$ ($k\left(s_L\right)\geq s_W$), then the leader works harder (as hard) as any individual worker.
Proof: $\tilde{e}(\theta)=k\left(s_L\right)\theta$, while $\hat{e}(\theta)=s_W\theta$. □
Claim 2: $k\left(s_L\right)>s_L$.
Proof: Since $s_L\in(0,1)$,

$$s_L^2<4s_L-3s_L^2.$$

Hence,

$$s_L<\frac{s_L+\sqrt{4s_L-3s_L^2}}{2}=k\left(s_L\right).\square$$

We begin by showing that the leader's share is declining in N: The cross-partial derivative of (18) with respect to s_L and N is

$$-\frac{1-s_L}{\left(N-1\right)^2}<0.$$

Since this is negative, the usual comparative statics entail that s_L is declining in N.

N	s_L	s_W	$k(s_L)$
2	.3333	.6667	.6667
3	.2149	.3926	.5320
4	.1818	.2727	.4871
5	.1668	.2083	.4655
6	.1584	.1683	.4529

Table 1: Data on Optimal Contracts and Efforts

The first-order condition for (18) is

$$[1 - k(s_L)]\, k'(s_L) - 1 + \frac{1 - s_L}{N - 1} = 0. \tag{34}$$

Let $N \to \infty$ and solve (34). This yields a lower bound for s_L of

$$\frac{10}{9} - \frac{19 \times 2^{\frac{1}{3}}}{9\left(187 + 9\sqrt{93}\right)^{\frac{1}{3}}} - \frac{\left(187 + 9\sqrt{93}\right)^{\frac{1}{3}}}{9 \times 2^{\frac{1}{3}}},$$

which is approximately .128843. Since

$$s_W = \frac{1 - s_L}{N - 1},$$

it follows from the lower bound that $s_L > s_W$ for $N \geq 7$. It then follows from Claims 1 and 2 that the lender works strictly harder than any worker for $N \geq 7$.

We need only consider $N \leq 6$ to complete the proof. Table 1 summarizes the relevant data. ■

Proof of Proposition 8: Working backwards, consider the last worker, N.[24] Assuming he owns the firm, he chooses e to maximize

$$\theta e - d(e) \tag{35}$$

regardless of the efforts supplied by the previous workers; that is, he supplies $e^{FB}(\theta)$. He supplies no effort if he does not take possession of the firm. Hence, the surplus created by selling him the firm is

$$\theta e^{FB}(\theta) - d\left[e^{FB}(\theta)\right].$$

[24] Here the more natural index of workers corresponds to the order in which they work, so this will be the indexing adopted for this proof.

Assume in bargaining with the $N-1$st worker that he captures $1-\beta_{N-1}$ of this positive surplus. Since the $N-1$st worker's share of the surplus is independent of his effort, he also chooses e to maximize (35) regardless of the efforts supplied by the previous workers. More generally, if the nth worker captures β_n of the surplus from selling to the $n+1$st worker, then, since $\beta_n, \ldots, \beta_{N-1}$ do not depend on e_n, nor do efforts e_m, $m = n+1, \ldots, N$, the nth worker chooses e to maximize (35) regardless of the efforts supplied by the previous workers. So by induction, all workers supply $e^{FB}(\theta)$. ■

Proof of Proposition 11: As before, let $e(\theta)$ be a worker's dominant strategy under an equal shares contract given he believes the state is θ. As a function of her announcement, $\hat{\theta}_i$, L_i's expected payoff is

$$\int_{\hat{\theta}_i}^{1} \left(ze(z) - d\left[e(z)\right]\right) f(z)\, dz + \int_0^{\hat{\theta}_i} \left(\frac{\theta_i}{N}\left[e(\theta_i) + (N-1)\, e\left(\hat{\theta}_i\right)\right] - d\left[e(\theta_i)\right]\right) f(z)\, dz,$$

where $f(\cdot)$ is the density over θ_j. The first-order condition is

$$\begin{aligned}
&-\left(\hat{\theta}_i e\left(\hat{\theta}_i\right) - d\left[e\left(\hat{\theta}_i\right)\right]\right) f\left(\hat{\theta}_i\right) + \left(\frac{\theta_i}{N}\left[e(\theta_i) + (N-1)\, e\left(\hat{\theta}_i\right)\right] - d\left[e(\theta_i)\right]\right) f\left(\hat{\theta}_i\right) \\
&+ \int_0^{\hat{\theta}_i} \frac{\theta_i}{N}(N-1)\, e'\left(\hat{\theta}_i\right) f(z)\, dz = 0. \qquad (36)
\end{aligned}$$

If $\hat{\theta}_i = \theta_i$, then the first line is zero, but the second line is positive. Hence, truth-telling is not a best response. The rest of the proposition follows straightforwardly. ■

Proof of Proposition 12: Given Proposition 1, all we need do is show this contract induces truth-telling by each leader. The first-order condition for maximizing a leader's expected utility is

$$\text{Left-hand side of (36)} + \partial t_{L_i}/\partial\hat{\theta}_i = 0,$$

where

$$\begin{aligned}
\frac{\partial t_{L_i}}{\partial \hat{\theta}_i} &= -\frac{\hat{\theta}_i}{N}(N-1)\, e'\left(\hat{\theta}_i\right) F\left(\hat{\theta}_i\right) \\
&= -\int_0^{\hat{\theta}_i} \frac{\hat{\theta}_i}{N}(N-1)\, e'\left(\hat{\theta}_i\right) f(z)\, dz.
\end{aligned}$$

Clearly, truth-telling solves the first-order condition. Moreover, simple algebra shows that the left-hand side of this first-order condition is positive for

$\hat{\theta}_i < \theta_i$ (recall $e(\cdot)$ is increasing) and negative for $\hat{\theta}_i > \theta_i$, so the first-order condition is sufficient as well as necessary. ■

Proof of Proposition 14: It is readily shown that this model satisfies the assumptions of Mailath (1987)'s Theorem 2, which in turn means $\tilde{e}'(\theta) > 0$. Comparing the first-order condition for a worker with share s_L,

$$s_L\theta_i - d'(e) = 0,$$

to (27) reveals that $\tilde{e}(\theta_i) > \hat{e}(s_L\theta_i)$ (recall $\tilde{e}'(\theta) > 0$). ■

References

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, **108**, 797–818.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural changes as informational cascades. *Journal of Political Economy*, **100**, 992–1026.

Caillaud, B. and Hermalin, B. E. (1993). The use of an agent in a signalling model. *Journal of Economic Theory*, **60**, 83–113.

Cho, I.-K. and Kreps, D. (1987). Signaling games and stable equilibria. *Quarterly Journal of Economics*, **102**, 179–222.

Demski, J. S. and Sappington, D. E. M. (1991). Resolving double moral hazard problems with buyout agreements. *RAND Journal of Economics*, **22**, 232–240.

Edlin, A. S. and Hermalin, B. E. (1996). Contract renegotiation in double moral hazard problems. Working Paper, Department of Economics, University of California at Berkeley.

Fudenberg, D. and Tirole, J. (1990). Moral hazard and renegotiation in agency contracts. *Econometrica*, **58**, 1279–1320.

Gal-Or, E. (1987). First mover disadvantages with private information. *The Review of Economic Studies*, **54**(2), 279–292. April.

Handy, C. (1993). *Understanding Organizations*. Oxford University Press, Oxford, England.

Hermalin, B. E. and Katz, M. L. (1991). Moral hazard and verifiability: The effects of renegotiation in agency. *Econometrica*, **59**, 1735–1753.

Hermalin, B. E. and Katz, M. L. (1993). Judicial modification of contracts between sophisticated parties: A more complete view of incomplete contracts and their breach. *The Journal of Law, Economics, & Organization*, **9**, 230–255.

Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, **13**, 324–340.

Kotter, J. P. (1996). *Leading Change*. Harvard Business School Press, Boston.

Levitt, S. D. and Snyder, C. M. (1996). Is no news bad news? Information transmission and the role of "early warning" in the principal-agent model. Working Paper, Harvard Society of Fellows.

Ma, C.-t. A. (1994). Renegotiation and optimality in agency contracts. *Review of Economic-Studies*, **61**, 109–129.

Mailath, G. J. (1987). Incentive compatibility in signaling games with a continuum of types. *Econometrica*, **55**, 1349–1365.

Mailath, G. J. (1993). Endogenous sequencing of firm decisions. *Journal of Economic Theory*, **59**(1), 169–182. February.

McGregor, D. (1966). *Leadership and Motivation*. MIT Press, Cambridge, MA.

Prendergast, C. (1993). A theory of "yes men". *The American Economic Review*, **83**, 757–770.

Rasmusen, E. (1987). Moral hazard in risk-averse teams. *RAND Journal of Economics*, **18**, 428–435.

Scharfstein, D. S. and Stein, J. C. (1990). Herd behavior and investment. *The American Economic Review*, **80**, 465–479.

Varian, H. R. (1994). Sequential contributions to public goods. *Journal of Public Economics*, **53**, 165–186.