Scapegoats and Transparency in Organizations

Björn Segendorff*†

June 5, 2000

Abstract

A separating equilibrium in which competent (incompetent) leaders choose competent (incompetent) co-workers is investigated. An outside observer rewards the leader at good policy outcomes. The incompetent co-worker can, at bad outcomes, be used as scapegoat. By assumption, the leader may fail in blaming the scapegoat. Two different assumptions on the outside observer's information set are compared. If she cannot observe a failed attempt, the separating equilibrium exists only if two non-mimicking constraints are met. If she can observe a failed attempt, an additional constraint is added due to the possibility of partial mimicking.

1. Introduction

A competent leader may choose an incompetent co-worker for different reasons. If it is legitimate for a leader to blame a subordinate there exists an insurance motive; if things go wrong he can blame the co-worker and thereby signal his own competence to an outside observer. However, such a scapegoat strategy may not always work as planned. The leader may, as an example, fail to find sufficient proofs of the co-worker's lack of competence - even though only relatively weak proofs may be required. Of importance for the value of the scapegoat strategy is also to what extent the outside observer has access to information. The objective of this study is to investigate under what circumstances an organizational culture in which competent (incompetent) leaders choose incompetent (competent) co-workers may exist because of the insurance motive when it is uncertain whether the leader will be able to blame the co-worker or not. In particular is the impact of two different informational assumptions studied; the outside observer may, or may not, be able to observe a failed attempt to blame the co-worker.

The leader, being either competent or incompetent, chooses the type of the co-worker; incompetent if the leader himself is competent and competent if he is incompetent. The types of the leader and the co-worker are private information to the leader (the incentives of the co-worker is abstracted from in order make the study as simple as possible). The leader and the co-worker undertake an activity (form a policy) with a random outcome but such that the expected outcome is increasing in their competences. The competent type of leader signals his

* Department of Economics, Stockholm School of Economics, P.O Box 6501, SE-113 83 STOCKHOLM and Department of Economics, Princeton University, Fisher Hall, Princeton, NJ 08544.

†I thank the IUI seminar participants for helpful comments. Financial support from *Jan Wallander and Tom Hedelius' Foundation* (J98/44) is gratefully acknowledged.

own ability by trying to blame his co-worker by proving him to be incompetent if the policy outcome has been unfavorable. The incompetent type of leader does not attempt do this. An attempt to blame the co-worker, i.e. to reveal his true type, succeeds with a given probability. The outside observer, henceforth the observer (she), can only observe the policy outcome and, if the co-worker has been blamed, eventually the co-worker's type. The observer's objective is to reward a competent leader and the objective of any type of leader is to achieve the reward. One natural interpretation of the situation described above is political organizations or institutions. The leader may be thought of as the president of a country or the leader of a political party. The co-worker could be the minister of finance, vice-president of the country or party or some member of the party's board depending on the interpretation. The outside observer can be thought of as the voters and the reward as being reelected or elected. The activity undertaken could be economic policy or an electoral campaign. It is worth noticing that the co-worker doesn't have to be a person, he could be a consulting firm or a group of people possibly outside the government or the party. The model may in some cases also apply to non-political organizations such as firms, public administration etc.

In the first setting investigated, the observer is assumed to observe both failed and successful attempts to blame the co-worker. This is the assumption of transparency and it has three interpretations. First, the process of finding proofs of the co-worker's true ability is easily observable to any outside observer. Secondly, the process of finding proofs may not be easily observable to any outside observer but the observer is in the privileged position of having access to additional information. Thirdly, the leader has some proofs of the co-worker's true ability but must decide whether to blame the co-worker or not before knowing whether the proofs are sufficient or not. In any of these three interpretations the observer would observe a failed attempt. In equilibrium, only the competent type of leader attempts to blame his co-worker at unfavorable policy outcomes. The leader is rewarded at favorable outcomes and at bad outcomes if he has attempted to blame his co-worker, successfully or not. A common result when studying separating equilibria is that non-mimicking is a sufficient condition for existence, i.e. the size of the reward or the leader's valuation of the reward should be such that it does not pay for any type of leader to mimic the behavior of the other type. Here, transparency adds one more constraint to the two non-mimicking constraints. The additional constraint says that the reward cannot be too high relatively to the probability with which an attempt to blame the co-worker is successful. This is because the incompetent type of leader may partially mimic the competent type by trying to blame his competent co-worker at policy outcomes below the threshold hoping to fail. If he fails, the observer cannot distinguish him from the competent type of leader and will consequently reward him. This deviation is profitable when the reward is too high relative to the probability of an attempt being successful.

In the second setting, the observer is not able to observe a failed attempt to blame the co-worker. This assumption is called non-transparency and it is at hand if (a) the proof-gathering process is not observable and (b) the leader knows whether the proofs are sufficient or not before making the decision whether to blame the co-worker or not. This might be the case if it is in the interest of the leader to secretly search for evidence and not making an accusation he cannot prove. As before, only the competent type of leader tries to blame his co-worker. The leader is rewarded at favorable policy outcomes and at unfavorable outcomes if he has been successful in blaming his co-worker. Partial mimicking is no longer an alternative for the incompetent type of leader and the separating equilibrium exists if the two non-mimicking constraints are met. Because the leader is only rewarded after a successful attempt to blame an incompetent co-worker the non-mimicking constraints turn out to be functions of the probability with which an attempt succeeds. The higher this probability is, the lower must the reward be. In the case where a separating equilibrium can exists under both transparency and non-transparency and under a weak monotonicity assumption on the observer's cost of rewarding the leader, it turns out that the equilibrium under non-transparency may be Pareto dominated by that under transparency. The opposite does not hold.

The most closely related study is Segendorff [13] who investigates a model closely related to the one below. He lets the leader succeed in blaming the co-worker with probability one but makes the co-workers participation voluntary resulting in two participation constraints that must be met in equilibrium additional to the non-mimicking constraints. His set-up also allows for situations where the leader and the co-worker compete for the reward. The present study is different since the leader may fail in blaming the co-worker and two different assumptions about the observer's information set are investigated.

There is a literature on decision makers' strategic use of information to signal ability. In Levy [8] well informed decision makers choose uninformed advisors in order to signal their own ability since contradicting one's advisor signals confidence in the own information. On the other hand, an uninformed decision maker is more in need of extra information and chooses a well informed advisor. Examples of other studies are Effinger and Polborn [3], Gibbons and Murphy [5], Jeon [7], Meyer and Vickers [9], and Trueman [14]. Leaders motivated by reputational concern have also been studied in the political economy literature. A leader may, prior to an election, strategically manipulate the policy in order to improve his reelection probability resulting in a political business cycle. In studies such as Persson and Tabellini [12], Rogoff [10], and Rogoff and Sibert [11] voters can distinguish competent leaders from incompetent by their choice of policy. In other studies, such as Besley and Case [1], Ferejohn [4], and Harrington [6], the policy is assumed to be unobservable and voters observe only the outcome. The leader chooses level of effort that positively affects the expected policy outcome and he is rewarded if the policy outcome is above some certain threshold.

The study below differs from the strategic information literature by (i) not focusing on strategic use of information but on the insurance motive and (ii) by treating the policy outcome as a function of the abilities of both the leader and the co-worker. It differs from the political economy literature by not trying to explain political business cycles but the impact of informational assumptions on the conditions ensuring the existence of a separating equilibrium. Moreover, it differs from Persson and Tabellini [12], Rogoff [10], and Rogoff and Sibert [11] in that the competent leader lowers the expected policy outcome when employing an incompetent co-worker and by assuming that the leader's action is unobservable. It differs from Besley and Case [1], Ferejohn [4], and Harrington [6] by assuming that the leader's choice of action may be verified if it is in the interest of the leader.

The basic model is presented in Section 2 where transparency is assumed. The assumption of transparency is replaced by non-transparency in Section 3. There is a short welfare analysis in Section 4. Section 5 contains a summary and an example. All proofs are given in the Appendix.

2. The Signalling Game

There are two players, the leader L (he) and the observer O (she). The leader may be of two types, competent (L=1) or incompetent (L=0), and p is the probability of him being competent. After observing his own type, L chooses the level of competence of his co-worker who is denoted C, competent (C=1) or incompetent (C=0). The types L and C are private information to L but p is common knowledge. With help from the co-worker the leader undertakes some activity with outcome

$$x = g(\mathsf{L}, \mathsf{C}) + \epsilon. \tag{2.1}$$

The random component ϵ has a continuous distribution over $\mathbb{R}$ and expected value $E[\epsilon] = 0$. Let $f(\epsilon)$ be its probability density function and $F(\epsilon)$ its cumulative density function. Here it is assumed that $f' \geq 0$ for all $\epsilon < 0$ and $f' \leq 0$ for all $\epsilon > 0$. Examples of such distributions are the normal, Cauchy, standard logistic, and t-distributions. The function g is increasing in the competence of L and C but at a decreasing rate, i.e. $g(0,0) = 0$, $g(0,1) = 1$, $g(1,0) = \alpha$, and $g(1,1) = \alpha + \gamma$ where α is the relative importance of L's competence for the policy outcome. It is natural to believe L's competence to be at least as important as C's competence and here it is assumed that $\alpha \geq 1$. The parameter $\gamma > 0$ is the competent co-worker's contribution to the expected policy outcome when working for the competent type of leader. Here, C's contribution may or may not decrease with the competence of L. After having observed the policy outcome, L has the opportunity to blame C, i.e. to reveal C's type to O at a cost $c \geq 0$. L can not lie and this is a reasonable assumption in those situations O requires L to proof his accusation and c can be thought of as the cost of searching for such proofs. Let $B = 0$ if L doesn't blame the co-worker and let $B = 1$ if he does. Assume further that L fails to provide a sufficient proof with probability $1 - q$ in which case O learns that there was a failed attempt to blame the co-worker. This assumption, that O learns also about failed attempts, corresponds to a situation where the organization is transparent. She has the ability to observe some, but not all, of the interaction between L and C. However, transparency is not always a good description of the world and in Section 3 this assumption is relaxed. Here, in the case of transparency, O observes the policy outcome and eventually the co-worker's type or a noise (n). Her information set is $\{x, \mathcal{C}\}$ where $\mathcal{C} \in \{\emptyset, n, 0, 1\}$ is O's observation of C. On the basis of her information, O updates her beliefs over L's competence. She cares about the leader's competence and wants to acknowledge/reward a competent leader but not an incompetent. The acknowledgment has many interpretations depending on the interpretation of the model but here it is thought of as the reelection of a politician. What is important is that a competent leader is more valuable to O if he is acknowledged than if not; A competent politician is more valuable in office than out of office. The observer also cares about the policy outcome and the cost of an eventual acknowledgment. Her utility is given by the vNM utility function

$$v_O(A) = A(\mathsf{L} - \theta) + p(1 - A) + \beta x \tag{2.2}$$

where $A = 1$ if she decides to acknowledge L and $A = 0$ otherwise. Her valuation of a competent (incompetent) leader is 1 (0) and the cost of an acknowledgment is $\theta \geq 0$. Assuming that the net value of acknowledge a competent leader is positive ($\theta < 1 - p$) her objective is to acknowledge L only if he is competent. If she chooses not to acknowledge L she receives p which can be interpreted as the probability of finding a competent leader somewhere else. This probability is then assumed to be equal to the unconditional probability of L being competent. Thus, P will only acknowledge L if she believes him to be competent with at least probability $p+\theta$. The parameter β expresses the relative weight she attaches to the policy outcome. L, on the other hand, desires an acknowledgment but has preferences also over the policy outcome. His utility is given by the vNM utility function

$$v_L(C,B) = x + A\phi - Bc \tag{2.3}$$

where ϕ is his valuation being acknowledged and is henceforth referred to as the reward. The reward and the cost of an acknowledgment is treated as parameters and will not be determined within the model.

The set-up described above is formalized by the game $G = (N, S, u)$ where $N = \{L, O\}$ is the set of players, S is the strategy sets, and u the payoff functions. A strategy for L is a mapping from his set of types and the set of policy outcomes to the set of types of co-worker and decisions whether to blame the co-worker or not, i.e. $s_L : \{0,1\} \times \mathbb{R} \twoheadrightarrow \{0,1\}^2$. His strategy set S_L is the set of all such mappings. Analogously, a strategy for O is a mapping from the set of possible observations to the set of acknowledgment decisions, $s_O : \{\emptyset, n, 0, 1\} \times \mathbb{R} \twoheadrightarrow \{0,1\}$, and her strategy set S_L is the set of all such mappings. Let $S = S_L \times S_O$ and let $s = (s_L, s_O)$ denote a strategy profile. The payoff functions u_L and u_P are mappings from the set of strategies to the real numbers, $u_i : S \to \mathbb{R}$ for $i = L, P$, i.e. $u_i(s)$ is the expected value of v_i when the strategy profile s is played. Let $u = u_L \times u_O$.

2.1. Separating Equilibrium in Transparent Organizations

Only separating equilibria are investigated below and by equilibrium is thus meant a separating equilibrium. The (separating) equilibrium of interest is such that L, if incompetent, chooses a competent co-worker and, if competent, chooses an incompetent co-worker. When the policy outcome is unfavorable in a sense explained below the competent type of leader blames his co-worker while the incompetent type never blames his co-worker. O acknowledges L if she observes a favorable policy outcome or if the policy outcome is unfavorable and an incompetent co-worker has been blamed or an attempt has been made.

A policy outcome is favorable for L if it makes O acknowledge him without L having to blame his co-worker, i.e. if O believes L to be competent with at least probability $p+\theta$ having observed only the outcome x but taking the separating choices of co-worker into account.

Definition 1. A policy outcome x is favorable if

$$\Pr(\mathsf{L}=1 \mid \mathsf{L} \neq \mathsf{C}) \equiv \frac{pf(x-a)}{pf(x-a) + (1-p) f(x-1)} \geq p + \theta$$

and the set of favorable policy outcomes is

$$X = \{x \mid \Pr(\mathsf{L}=1 \mid \mathsf{L} \neq \mathsf{C}) \geq p\}.$$

A policy outcome is unfavorable if it is not favorable.

For the assumed class of distributions X is an interval $[\overline{x}, +\infty)$ and a policy outcome is favorable if it is above the threshold $\overline{x}$. In the following it is assumed that X is non-empty, i.e. that the net value of acknowledging a competent leader is sufficiently high.

The two types of leader differ in their choices of co-worker and set of policy outcomes at which they blame their co-worker (X for the competent type of leader and $\emptyset$ for the incompetent type). The suggested strategy of the leader is thus a pair of step functions, one for each type

$$s_L^S(\mathsf{L}, x) = \begin{cases} \text{If } \mathsf{L}=0 \text{ then } \mathsf{C}=1 \text{ and } B=0 \text{ for all } x \\ \text{If } \mathsf{L}=1 \text{ then } \mathsf{C}=0,\ B=0 \text{ for all } x \notin X, \text{ and } B=1 \text{ for all } x \in X \end{cases} \tag{2.4}$$

The strategy of the observer is simply to acknowledge L if she believes him to be competent with at least probability p. Let $\pi(\mathcal{C}, x) = \Pr\left(\mathsf{L}=1 \mid \{\mathcal{C}, x\}, s_L^S\right)$ be O's belief over L's type when she observes $\{\mathcal{C}, x\}$ and she beliefs L's strategy to be s_L^S. Then,

$$s_O^S(\mathcal{C}, x) = \{A = 1 \text{ if } \pi(\mathcal{C}, x) \geq p + \theta \text{ and } A = 0 \text{ otherwise}\,. \tag{2.5}$$

From L's strategy it follows that $\pi(\mathcal{C}, x)$ is well defined for all information sets $\{\mathcal{C}, x\} \in \{\emptyset\} \times \mathbb{R}$ and all informations sets $\{\mathcal{C}, x\} \in \{0, n\} \times \mathbb{R}\backslash X$ where $\mathbb{R}\backslash X$ denotes the complement to X. However, $\pi(\mathcal{C}, x)$ is not well defined for out-of-equilibrium observations and there are two types of out-of-equilibrium observations; she may observe a competent co-worker being blamed or she may observe a noise or an incompetent co-worker being blamed at a favorable policy outcome. Since Bayes' rule does not apply to out-of-equilibrium events any beliefs can be assigned to these observations. The formation of out-of-equilibrium beliefs is assumed to have the following three reasonable properties; (a) O should form her beliefs in a consistent way for all out-of-equilibrium observations, (b) O should use only available and verifiable information, and (c) the belief formation should not be conflicting with the intuitive criterion (Cho-Kreps [2]).

Assumption 1. For all observations $\{\mathcal{C}, x\} \in \{0, n\} \times X \cup \{1\} \times \mathbb{R}$ the observer's out-of-equilibrium beliefs are given by

$$\pi(\mathcal{C}, x) = \frac{pf(x - g(1, \mathcal{C}))}{pf(x - g(1, \mathcal{C})) + (1-p) f(x - g(0, \mathcal{C}))} \tag{2.6}$$

for $\mathcal{C} \in \{0, 1\}$ and $\pi(n, x) = \lambda \pi(0, x) + (1 - \lambda)\pi(1, x)$ for some $0 \leq \lambda \leq 1$.

It is important to notice that this way of forming of out-of-equilibrium beliefs doesn't always make O believe L to be incompetent after having made an out-of-equilibrium observation. For some favorable policy outcomes she may believe L to be competent. However, it makes it unprofitable for L to deviate from s_L^S by blaming any type of co-worker at favorable policy outcomes since this can not increase his chances of being acknowledged, only decrease them. In the case O observes that a competent co-worker is blamed at an unfavorable policy outcome she doesn't believe the leader to be competent with high enough probability to acknowledge him.

A necessary condition for $s^S = \left(s_L^S, s_O^S\right)$ is that it should not be profitable for one type of leader to mimic the behavior of the other type, i.e. to play

$$s_S^M(\mathsf{L}, x) = \begin{cases} \text{If } \mathsf{L}=0 \text{ then } \mathsf{C}=0,\ B=0 \text{ for all } x \notin X, \text{ and } B=1 \text{ for all } x \in X \\ \text{If } \mathsf{L}=1, \text{ then } \mathsf{C}=1 \text{ and } B=0 \text{ for all } x. \end{cases} \tag{2.7}$$

Let E_{LC} be the set of realizations of ϵ such that the policy outcome is favorable when L is of type L and C is of type C, i.e.

$$E_{\mathsf{LC}} = \{\epsilon \mid \epsilon = x - g(\mathsf{L}, \mathsf{C}), x \in X\}\,.$$

Let the probability (cumulative density) of realizations of ϵ in E_{LC} be denoted $F(E_{\mathsf{LC}})$ which is the probability of a favorable policy outcome. Solving $u_L(s^S) \geq u_L(s_L^M, s_O^S)$ for the incompetent and competent types of leader, respectively, gives

$$\phi \leq \overline{\phi} \equiv \frac{(1 - F(E_{00}))\,c + 1}{1 - F(E_{01})} \tag{2.8}$$

and

$$\phi \geq \underline{\phi} \equiv \frac{(1 - F(E_{10}))\,c + \gamma}{1 - F(E_{11})}. \tag{2.9}$$

If the two inequalities above are satisfied simultaneously then no type of leader has an incentive to imitate the behavior of the other type. Appointing an incompetent co-worker lowers the expected policy outcome and thereby the probability of a favorable policy outcome. On the other hand, it enables L to be acknowledged at unfavorable policy outcomes: he gives up $(F(E_{\mathsf{L}1}) - F(E_{\mathsf{L}0}))\,\phi$ in order to gain $(1 - F(E_{\mathsf{L}0}))(\phi - c)$. In addition, he also gives up some expected utility, 1 or γ, derived directly from the policy outcome. In the separating equilibrium, this trade off must be profitable for the competent type of leader and unprofitable for the incompetent type. An eventual increase in c must thus be compensated by an increase in ϕ in order to maintain the original incentives. More precisely, the reward must increase by the increase in c times $(1 - F(E_{\mathsf{L}0}))\,/\,(1 - F(E_{\mathsf{L}1}))$ which is the relative probability of an unfavorable policy outcome when having an incompetent co-worker relative to when having a competent co-worker. It is not possible to say which for which type of leader this relative probability is greatest and, hence, not possible to say which of $\underline{\phi}$ and $\overline{\phi}$ that is the largest, this depends on c and the underlying distribution f. Comparing the expressions for $\underline{\phi}$ and $\overline{\phi}$ gives that $\overline{\phi}$ increases faster in c than what $\underline{\phi}$ does if

$$\frac{1 - F(E_{00})}{1 - F(E_{01})} - \frac{1 - F(E_{10})}{1 - F(E_{11})} \geq 0 \tag{2.10}$$

and $\overline{\phi} \geq \underline{\phi}$ if c is sufficiently large. Moreover, if the relative importance of C's ability decreases sufficiently with L's ability, i.e.

$$\gamma \leq \overline{\gamma} = \frac{1 - F(E_{11})}{1 - F(E_{01})} \tag{2.11}$$

then $\overline{\phi} \geq \underline{\phi}$ for some c. In the case $\overline{\phi} = \underline{\phi}$ for some $c \geq 0$, denote this cost $\underline{c}$.

Lemma 1. $\overline{\phi} \geq \underline{\phi}$ if either
(i) 2.10 and 2.11 hold,
(ii) 2.10 holds strictly and $c \geq \underline{c}$, or
(iii) 2.11 holds strictly and $c \leq \underline{c}$.

Proof. See Appendix.

No one of the two equations 2.10 and 2.11 is by itself a necessary or sufficient condition for the possibility to simultaneously satisfy the two non-mimicking constraints. In other words this does not critically depend on the properties of the distribution of ϵ or the size of the competent co-worker's contribution to the expected policy outcome. It is the combination of these two factors and the cost of blaming that determines the magnitudes of the constraints. For instance, Eq. 2.11 states that the competent co-worker's contribution should decrease with the competence of the leader, $\gamma \leq \underline{\gamma} < 1$, but according to Lemma 1(iii) the non-mimicking constraints may be simultaneously satisfied for some distributions f and costs c even if his contribution increases with the competence of the leader, i.e. $\gamma > 1$.

Verifying that mimicking is non-profitable is usually sufficient for the existence of a separating equilibrium but not in this case because of the possibility of partial mimicking where the incompetent type of leader tries to blame his competent co-worker at unfavorable policy outcomes. With probability q C's type is revealed and due to P's out-of-equilibrium beliefs he will not be acknowledged. However, with probability $1 - q$ P receives the noisy signal n and is incompetent to distinguish the incompetent type of leader from the competent and acknowledges him. This deviation must be non-profitable in equilibrium and the probability of revealing the co-worker's type can therefore not be too low

$$q \geq \frac{\phi - c}{\phi} \tag{2.12}$$

The expected payoff from partial mimicking increases with the reward and a high reward must be balanced with a low probability of transmitting the noisy signal. Partial mimicking would otherwise be profitable and s^S wouldn't be an equilibrium. Moreover, there can not exist a separating equilibrium in which both types of leader blame their co-workers at unfavorable policy outcomes. O would, by Bayes' rule, believe L to be incompetent after having observed a failed attempt at a bad policy and, hence, no type of leader would have an incentive to appoint an incompetent co-worker and there would be no separating equilibrium.

The proposed strategy profile s^S is an equilibrium if both mimicking and partial mimicking are non-profitable. Proposition 1 states that this is the case if at least one of the conditions in Lemma 1 holds and the probability of revealing C's type is sufficiently high.

Proposition 1. If (i), (ii) or (iii) in Lemma 1 holds and $q \geq \left(\underline{\phi} - c\right)/\underline{\phi}$ then s^S constitutes a NE (Nash equilibrium) for some reward ϕ.

Proof. See Appendix.

It is easy to see why Proposition 1 holds. Let (i), (ii) or (iii) in Lemma 1 hold. Then $\overline{\phi} \geq \underline{\phi}$ and letting $\phi = \underline{\phi}$ makes mimicking non-profitable for both types of leader. Partial mimicking is non-profitable for the incompetent type since $q \geq \left(\underline{\phi} - c\right)/\underline{\phi}$. This makes s^S a Nash equilibrium and Corollary 1 characterizes the interval of rewards that are consistent with equilibrium.

Corollary 1. Suppose (i), (ii) or (iii) in Lemma 1 holds. Then, if
(i) $\left(\underline{\phi} - c\right)/\underline{\phi} \leq q < \left(\overline{\phi} - c\right)/\overline{\phi}$ then s^S is a NE for all $\phi \in \left[\underline{\phi}, \frac{c}{1-q}\right]$ and if
(ii) $q \geq \left(\overline{\phi} - c\right)/\overline{\phi}$ then s^S is a NE for all $\phi \in \left[\underline{\phi}, \overline{\phi}\right]$.

Proof. See Appendix.

The interval of rewards consistent with equilibrium is restricted so that no kind of mimicking is profitable within the specified interval of rewards. It is restricted from below by the competent

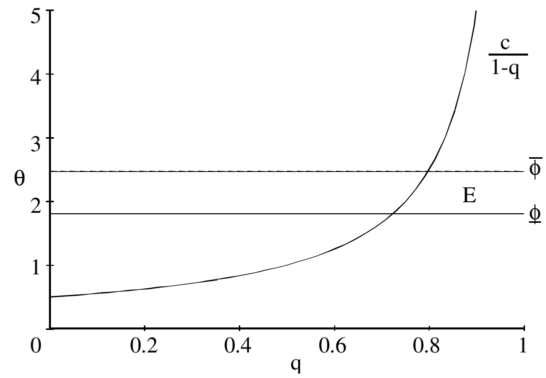

Figure 2.1: The separating equilibrium exists for all rewards and probabilities q in the equilibrium zone E.

type's non-mimicking constraint and from above it is restricted by either the incompetent type's non-partial mimicking constraint or his non-mimicking constraint. Which of the two latter constraints that will be restricting the interval of rewards depends on q, the probability with which the co-worker's type is conveyed, see Figure 2.1.

One important consequence of transparency is that it enables the incompetent type of leader to try to deceive O by partial mimicking. This is possible only because in equilibrium O must believe L to be competent whenever she observes a failed attempt to blame the co-worker at unfavorable policy outcomes. The alternative, that she believes L to be incompetent, can not be supported in equilibrium; suppose O acknowledged L only if she observed a successful attempt to blame an incompetent co-worker. Partial mimicking is then not an alternative for the incompetent type of leader and only the competent type of leader will try to blame his co-worker. This means, by Bayes' rule, that O must believe L to be competent whenever she observes a failed attempt. She will consequently acknowledge L which contradicts the assumption that she would not.

The assumption of transparency is not always an appropriate assumption because the observer may not have the ability to observe and in the next section the transparency assumption is exchanged to a non-transparency assumption.

3. Non-Transparent Organizations

Assume, as before, that L fails to blame the co-worker with probability $1 - q$. O learns only about successful attempts and is therefore ignorant about an eventual failed attempt, i.e. $\mathcal{C} \in \{\emptyset, 0, 1\}$ where $\mathcal{C} = \emptyset$ when L has not blamed the co-worker or when such an attempt has been unsuccessful. This is the assumption of non-transparency. There are many plausible reasons for non-transparency. L and O might be separated geographically so that O can not monitor L or O may not have the knowledge to distinguish a proof-gathering activity from the usual daily activity. Alternatively, L may secretly gather proofs that he puts forward only if he knows that they are sufficient. This may be the case when there is a large cost connected with failing an attempt publicly.

Let $G' = (N, S', u')$ be the signalling game under the assumption of non-transparency. L's strategy set S_L is as before but O's set of strategies is different. A strategy for O is a mapping $s_O : \{\emptyset, 0, 1\} \times \mathbb{R} \twoheadrightarrow \{0, 1\}$ and S'_O is the set of all such mappings. Let $S' = S_L \times S'_O$ be the combined strategy set. The payoff functions u'_L and u'_O are mappings from the set of strategies to the real numbers, $u'_i : S' \to \mathbb{R}$ for $i = L, O$, and $u' = u'_L \times u'_O$. The suggested equilibrium strategy profile is the same as in Section 2, and the equilibrium are defined by Bayes' rule as before. Also the set of favorable policy outcomes is the same as in Section 2. However, the out-of-equilibrium beliefs (Assumption 1) must be redefined to be over the new domain $\{0\} \times X \cup \{1\} \times \mathbb{R}$ which differ from the previous because O can not observe the signal n. Now any out-of-equilibrium observation involves O observing the co-worker's type and the out-of-equilibrium beliefs are as given by 2.6 in Assumption 1 with the exception that the signal n no longer can be observed.

Partial mimicking is not possible in a non-transparent organization; suppose the incompetent type partially mimics the competent type as described in Section 2. If he fails in blaming the co-worker then O doesn't learn anything about the attempt and believing that no attempt has been made she concludes that L is incompetent. If L succeeds, O believes him to be incompetent and doesn't acknowledge him. Partial mimicking is consequently a non-profitable deviation that doesn't need to be considered. and for s^S to constitute a Nash equilibrium it is sufficient that mimicking is non-profitable, i.e. $u_L\left(s^S\right) \geq u_L\left(s_L^M, s_O^S\right)$ for $\mathsf{L} = 0, 1$. Solving the inequality for the incompetent type gives

$$\phi \leq \overline{\phi}^N(q) = \begin{cases} \overline{\phi}\frac{1-F(E_{01})}{q(1-F(E_{00}))-(F(E_{01})-F(E_{00}))} & if \quad q \in (\overline{q}, 1] \\ +\infty & if \quad q \in [0, \overline{q}] \end{cases} \tag{3.1}$$

where $\overline{q} = \left(F\left(E_{01}\right) - F\left(E_{00}\right)\right) / \left(1 - F\left(E_{00}\right)\right)$. For the competent type

$$\phi \geq \underline{\phi}^N(q) = \begin{cases} \underline{\phi}\frac{1-F(E_{11})}{q(1-F(E_{10}))-(F(E_{11})-F(E_{10}))} & if \quad q \in \left(\underline{q}, 1\right] \\ +\infty & if \quad q \in \left[0, \underline{q}\right] \end{cases} \tag{3.2}$$

where $\underline{q} = \left(F\left(E_{11}\right) - F\left(E_{10}\right)\right) / \left(1 - F\left(E_{10}\right)\right)$. The two piece-wise defined functions above deserve some explanation.

The functions $\underline{\phi}^N$ and $\overline{\phi}^N$ specifies the lower- and upper bounds on the interval of rewards that are consistent with equilibrium, just as $\underline{\phi}$ and $\overline{\phi}$ did in Section 2. A high (low) probability of a successful blaming attempt gives a high (low) expected payoff from having an incompetent co-worker while the expected payoff from having a competent co-worker is the same in both cases. The expected payoff from having an incompetent co-worker increases faster in the reward than what the expected payoff from having a competent co-worker does. Starting in a situation where L is indifferent between appointing a competent or an incompetent co-worker, a decrease in the success probability q must be compensated by an increase in the reward in order to restore indifference. The function $\underline{\phi}^N$ and $(\overline{\phi}^N)$ is therefore decreasing in q for $q > \underline{q}$ $(q > \overline{q})$ goes to infinity as q approaches $\underline{q}$ and $(\overline{q})$ from above. Notice that $\underline{\phi}^N(1) = \underline{\phi}$ and $\overline{\phi}^N(1) = \overline{\phi}$.

There is an important difference between the two critical values $\underline{q}$ and $\overline{q}$. For the competent leader, when $q > \underline{q}$ there exist a reward $\underline{\phi}^N(q)$ such that he prefers the separating strategy to mimicking for all rewards larger than that reward. At $\underline{q}$ there no longer exist any such reward and the competent type strictly prefers mimicking to the separating strategy for all rewards. The latter is true for all $q \leq \underline{q}$ and $\underline{\phi}^N$ is thus equal to infinity for all such q. For the incompetent type the opposite is true. When $q > \overline{q}$ there exist a reward $\overline{\phi}^N$ such that he strictly prefers mimicking to the suggested equilibrium strategy for all rewards that are larger than that reward. At $\overline{q}$ there no longer exist such rewards and he strictly prefers the separating strategy to mimicking which is also true for all $q < \overline{q}$. Hence, $\overline{\phi}^N$ is equal to infinity for all $q \leq \overline{q}$. Two observations can be made. First, s^S can not be an equilibrium if $q \leq \underline{q}$ since the competent type would always deviate. Secondly, if 2.10 holds then $\overline{q} \geq \underline{q}$ and otherwise $\underline{q} \geq \overline{q}$. Moreover, $\underline{\phi}^N$ and $\overline{\phi}^N$ has a single-crossing property and investigating this property further gives three results. First, if $\overline{q} \geq \underline{q}$ and $\overline{\phi} \geq \underline{\phi}$, or if $\underline{q} > \overline{q}$ and $\underline{\phi} > \overline{\phi}$, then $\overline{\phi}^N$ does not cross $\underline{\phi}^N$ for any $q \geq \underline{q}$. Secondly, if $\underline{q} > \overline{q}$ and $\overline{\phi} > \underline{\phi}$, or $\overline{q} > \underline{q}$ and $\underline{\phi} > \overline{\phi}$ then there exists a unique $q^* \in \left(\underline{q}, 1\right]$ at which $\underline{\phi}^N$ crosses $\overline{\phi}^N$. Combining the single-crossing property with Lemma 1 shows that 3.2 and 3.1 can be satisfied simultaneously under some will-specified circumstances, see also Figure 3.1.

Lemma 2. $\overline{\phi}^N(q) \geq \underline{\phi}^N(q)$ if any of the following is true,
(i) 2.10 and 2.11 hold and $q \geq \underline{q}$,
(ii) 2.10 holds strictly, $c \geq \underline{c}$ and $q \geq \underline{q}$,
(iii) 2.10 holds strictly, $c \leq \underline{c}$ and $\underline{q} \leq q \leq q^*$, or
(iv) 2.11 holds strictly, $c \leq \underline{c}$ and $q \geq q^*$.

Proof. See Appendix.

If any of the four cases in Lemma 2 is at hand, then there exists at least one reward such that no type of leader has an incentive to deviate. More precisely, s^S is a Nash equilibrium for every reward that satisfies both 3.2 and 3.1.

Proposition 2. If (i), (ii), (iii) or (iv) in Lemma 2 holds, then s^S is a NE for all rewards $\phi \in \left[\underline{\phi}^N(q), \overline{\phi}^N(q)\right]$.

Proof. See Appendix.

The suggested strategy profile s^S is a Nash equilibrium if the reward ϕ is such that mimicking is non-profitable for both types of leader. Whether this is the case or not depends on the properties of f and g and the sizes of c and q. When f and g are such that 2.10 or 2.11 holds and c and q assumes appropriate values, the incentive constraints of the two types of leader can be mutually satisfied, i.e. $\overline{\phi}^N(q) \geq \underline{\phi}^N(q)$. Any reward between the upper- and lower bound satisfies the non-mimicking constraints making s^S an equilibrium.

4. Welfare Analysis

Assume that the parameters of the model, except for O's cost and L's valuation of an acknowledgment, are such that the separating equilibrium can exist both under transparency and

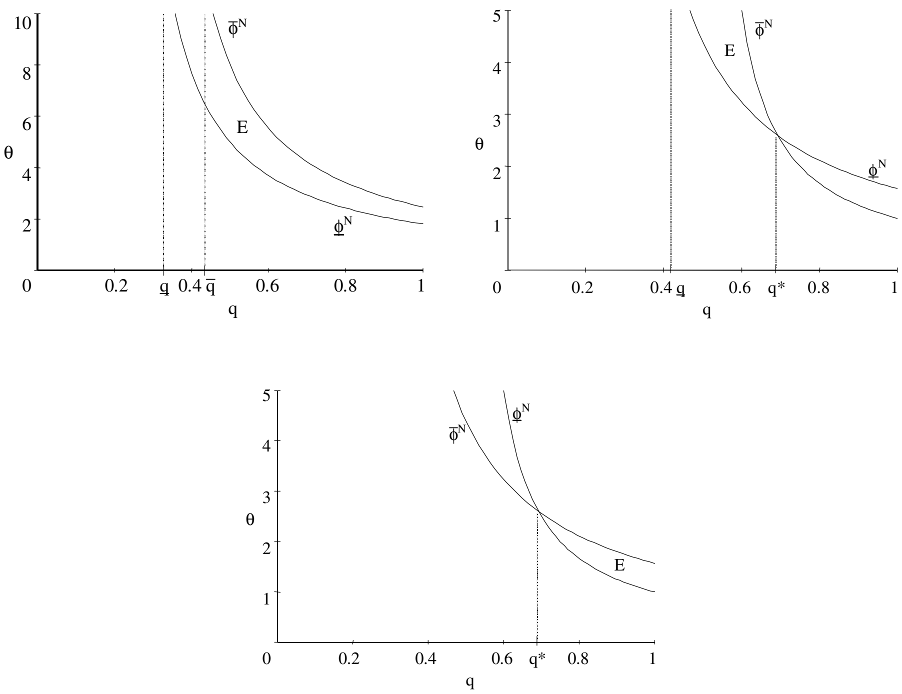

Figure 3.1: The equilibrium zone of rewards and probabilities q can have three different shapes. The upper left diagram illustrates the a situation where Lemma 2(i) or (ii) holds. The upper right diagram show the zone when Lemma 2(iii) holds and the diagram below when Lemma 2(iv) holds.

non-transparency. Let (ϕ, θ) and (ϕ', θ') be the valuation and cost of an acknowledgment associated with equilibrium under transparency and non-transparency, respectively. Comparing the expected utilities of O and the competent type of leader in both settings gives one condition for each player specifying when one setting is preferred to the other by that player. The incompetent type of leader is acknowledged witht he same probability in both settings and prefer the setting that has the highest reward.

Lemma 3. The expected utility of the observer is higher under transparency than non-transparency if and only if

$$\theta - \theta' \leq \frac{p(1-q)(1-F(E_{10}))(1-\theta'-p)}{p+(1-p)F(E_{01})} \tag{4.1}$$

and the corresponding is true for the competent type of leader if and only if

$$\phi' - \phi \leq (1-q)(1-F(E_{10}))\phi' \tag{4.2}$$

Proof. See Appendix.

Notice that the right-hand side of 4.1 is strictly positive since $1-\theta'-p>0$ is a necessary condition for X to be non-empty as assumed. O's welfare is thus higher under non-transparency only if the cost of acknowledging L is lower than under transparency. This is because the competent type of leader is acknowledged with a higher probability under transparency (probability 1) than under non-transparency. The lower probability under non-transparency lowers the expected payoff of both the observer and the competent leader, ceteris paribus, since the observer receives the gain $1-\theta$ with lower probability just as the competent leader receives his reward with lower probability. Keeping their expected payoffs equal in both settings requires the net gain and the reward to be higher under non-transparency than under transparency, i.e. $\theta' < \theta$ and $\phi > \phi'$. Since it is only the informational assumption that differs between the two settings the observer's cost of a particular reward should be the same in both settings. It is also natural to impose a weak monotonicity assumption stating that a large reward costs at least as much as a smaller reward.

Assumption 2. (Weak Monotonicity) In any setting, or between the two settings, $\phi' > \phi$ implies $\theta' \geq \theta$.

A consequence of Weak Monotonicity is that the two inequalities in Lemma 3 may be satisfied but not violated simultaneously.

Proposition 3. Under the assumption of Weak Monotonicity the separating equilibrium under transparency may Pareto dominate the separating equilibrium under non-transparency while the reversed is not true.

Proof. See Appendix

For the equilibrium under transparency to Pareto dominate that under non-transparency it is necessary that the reward in the former isn't lower than that in in the latter because of the welfare of the incompetent leader. Any equilibrium reward must satisfy the non-mimicking constraints, and in particular $\phi \leq \overline{\phi}$ and $\phi' \geq \underline{\phi}^N(q)$ where the latter constraint goes from $\underline{\phi}$ to $+\infty$ as q goes from 1 to $\underline{q}$. Hence, for q close enough to 1 there exists rewards ϕ and ϕ' consistent with equilibrium such that transparency Pareto dominates non-transparency. The intuition is that the impact of the informational assumptions decreases with q and vanishes as q goes to 1. For lower values of q the difference is greater and for sufficiently low values of q is the distance between the two constraints $\overline{\phi}$ and $\underline{\phi}^N(q)$ large enough to make all possible combinations of rewards and costs, (ϕ, θ) and (ϕ', θ'), violate either 4.1 or 4.2 given the monotonicity assumption. In that case no reward and cost structure in one setting can Pareto dominate a structure in the other setting.

5. Summary

The impact of two different informational assumptions, transparency and non-transparency, on the conditions for existence of a separating equilibrium has been investigated. As is usual in this kind of models, the separating equilibrium exists when it is not profitable for one type of leader to mimic the behavior of the other type. Under the assumption of non-transparency, the leader is not acknowledged if his attempt to blame the co-worker fails and the non-mimicking constraints thus *changes* with the probability of the leader's attempt to blame the co-worker being successful; the lower probability of successfully blaming the co-worker the higher must the leader's reward be when acknowledged. In the case of transparency things are different. The observer learns about failed attempts to blame the co-worker and rewards the leader also after a failed attempt. The competent leader will consequently always be rewarded in equilibrium. It follows that the non-mimicking constraints are *constant* with respect to probability of successfully blaming the con-worker. The higher probability of the competent type of leader to be acknowledged also makes the non-mimicking constraints assume *lower values* under transparency than under non-transparency. On the other hand, transparency opens up for the possibility of partial mimicking which *adds one more constraint* to the two non-mimicking constraints. The main findings are summarized in Table I.

Table I

	Transparency	Non-Transparency
Constraints	$\phi \geq \underline{\phi}$ $\phi \leq \overline{\phi}$ $\phi \leq c/\left(1-q\right)$	$\phi \geq \underline{\phi}^{N}\left(q\right)$ $\phi \leq \overline{\phi}^{N}$
Probability of $\mathsf{L}=0$ being acknowledged	$F\left(E_{01}\right)$	$F\left(E_{01}\right)$
Probabilility of $\mathsf{L}=1$ being acknowledged	1	$F\left(E_{10}\right)\left(1-q\right)+q$
Equilibrium value of q	$q \geq \left(\underline{\phi}-c\right)/\underline{\phi}$ $q \leq \left(\overline{\phi}-c\right)/\overline{\phi}$	$q \geq \underline{q}$ or $\underline{q} \leq q \leq q^{*}$ or $q \geq q^{*}$
Relative size of reward	$\underline{\phi} \leq \underline{\phi}^{N}\left(q\right)$ $\overline{\phi} \leq \overline{\phi}^{N}\left(q\right)$	
Pareto domination	Eventually > but never <	

6. A Numerical Example

The standard logistic distribution closely approximates the standard normal distribution. Let ϵ be standard logistically distributed, $F\left(\epsilon\right)=e^{\epsilon}/\left(1+e^{\epsilon}\right)$ and let L be competent with probability $p=0.5$. The relative contribution of his ability to the policy outcome is $\alpha=1.5$ and when working for the competent leader the competent co-worker contributes $\gamma=0.4$. The cost of blaming the co-worker is $c=0.5$ and the intensity of P's preferences over the policy outcome is $\beta=1$. These parameters are held constant throughout the example and they are used as a common starting point for the investigation of two different cases, $\left(\theta,\phi,q\right)=\left(0,2.5,0.9\right)$ and $\left(\theta',\phi',q'\right)=\left(0,2.5,0.9\right)$, using both informational assumptions. The example is described in Table II.

The case $\left(\theta,\phi,q\right)$ is the most favorable to both the observer and the leader. The observer is, without any costs, able to give the presumably competent leader a high reward. Under transparency it turns out that both 2.10 and 2.11 holds and $\overline{\phi} \geq \underline{\phi}$ for all $c \geq 0$ by Lemma 1(i). The reward $\theta=2.5$ satisfies both non-mimicking constraints. Moreover, q is sufficiently high given θ and satisfies the partial non-mimicking constraint. Also under non-transparency is q sufficiently high ($q>\underline{q}$, Lemma 2(i)) and θ satisfies the non-mimicking constraints. Hence, the separating equilibrium exists in both settings and here non-transparency is Pareto dominated by transparency.

The second case $\left(\theta',\phi',q'\right)$ is less favorable than the first, it is costly for O to acknowledge L who in turn valuates the acknowledgment less than in the first case. For these parameter values it turns out that only 2.11. It follows, by Lemma 1(iii), that c cannot be too high which is not the case here. Under the assumption of transparency the reward satisfies both mimicking constraints and q the partial non-mimicking constraint. The non-mimicking constraints are satisfied also under the assumption of non-transparency. In the latter setting q is also higher than required by Lemma 2(iii). The separating equilibrium exists in both settings and also here is non-transparency Pareto dominated by transparency.

Table II

$p = 0.5, \alpha = 1.5, \beta = 1, \gamma = 0.4, c = 0.5$

	Transparency		Non-Transparency
		2.10 *holds*	
		2.11 *holds*	
		$X = [1.125, +\infty)$	
$\theta = 0$		$\overline{\gamma} = 0.55$	
$\phi = 2.5$	$q \geq 0.8$		$\underline{q} = 0.2166$
$q = 0.9$	$\underline{\phi} = 1.8$		$\underline{\phi}^N (0.9) = 2.07$
	$\overline{\overline{\phi}} = 3.78$		$\overline{\overline{\phi}}^N (0.9) = 2.87$
	$u_O \left(s^S\right) = 1.89$		$u'_O \left(s^S\right) = 1.88$
	$u_{L=1} \left(s^S\right) = 3.78$		$u'_{L=1} \left(s^S\right) = 3.69$
		2.10 *doesn't hold*	
		2.11 *holds*	
		$X = [3.518, +\infty)$	
$\theta' = 0.1$		$\overline{\gamma} = 0.9$	
$\phi' = 1.4$		$\underline{c} = 73.24$	
$q' = 0.75$	$q \geq 0.64$		$q^* = 0.07$
	$\underline{\phi} = 1.01$		$\underline{\phi}^N (0.75) = 1.88$
	$\overline{\overline{\phi}} = 1.61$		$\overline{\overline{\phi}}^N (0.75) = 2.26$
	$u_O \left(s^S\right) = 1.92$		$u'_O \left(s^S\right) = 1.88$
	$u_{L=1} \left(s^S\right) = 2.48$		$u'_{L=1} \left(s^S\right) = 2.26$

Proof of Lemma 1. Rewriting $\overline{\phi} \geq \underline{\phi}$ gives

$$\left(\frac{1-F(E_{00})}{1-F(E_{01})}-\frac{1-F(E_{10})}{1-F(E_{11})}\right)c \geq \frac{\gamma}{1-F(E_{11})}-\frac{1}{1-F(E_{01})}. \tag{.1}$$

(i) Suppose 2.11 and 2.10 hold. The right-hand side of .1 is non-positive and the left-hand side is non-negative since c is non-negative by assumption. Hence, .1 holds for all $c \geq 0$. (ii) Suppose 2.10 holds strictly and 2.11 doesn't hold. Both sides of .1 is strictly positive and .1 holds for all $c \geq \underline{c}$ where

$$\underline{c} \equiv \frac{\frac{\gamma}{1-F(E_{11})}-\frac{1}{1-F(E_{01})}}{\frac{1-F(E_{00})}{1-F(E_{01})}-\frac{1-F(E_{10})}{1-F(E_{11})}} \geq 0. \tag{.2}$$

(iii) Suppose 2.11 holds strictly while 2.10 doesn't hold. Both sides of .1 are strictly negative since $c \geq 0$ by assumption. Hence, if $c \leq \underline{c}$ then .1 holds.

Proof of Proposition 1. $\overline{\phi} \geq \overline{\phi}$ since (i), (ii) or (iii) in Lemma 1 holds by assumption and $q \geq \left(\underline{\phi}-c\right)/\underline{\phi}$. Let $\phi = \underline{\phi}$. Consider any possible strategy $s_L' \neq s_L^S$ and assume that s_L' prescribes L to chose an incompetent leader. Then $s_L \succ_{\mathsf{L}=0} s_L^M$ by assumption and $s_L^M \succeq_{\mathsf{L}=0} s_L'$ by the properties of O's belief formation. The latter also gives $s_L^S \succeq_{\mathsf{L}=1} s_L'$. A deviation prescribing a incompetent co-worker can not be strictly profitable. Similarly, if s_L' prescribe L to chose a competent co-worker then $s_L^S \sim_{\mathsf{L}=1} s_L^M$ by assumption and $s_L^M \sim_{\mathsf{L}=1} s_L'$ by the properties of O's belief formation which also gives $s_L^S \succeq_{\mathsf{L}=0} s_L'$. Hence, there exists no strictly profitable deviation for any type of leader given the assumptions made above. s_O^S maximizes, by definition, O's expected utility. Hence, s^S is a NE for $\phi = \underline{\phi}$. ■

Proof of Corollary 1. $\overline{\phi} \geq \overline{\phi}$ since (i), (ii) or (iii) in Lemma 1 holds by assumption. From the proof of Proposition 1 it follows that if mimicking is a non-profitable deviation, then is s^S a NE. Hence, if $\left(\underline{\phi}-c\right)/\underline{\phi} \leq q < \left(\overline{\phi}-c\right)/\overline{\phi}$ then $\underline{\phi} \leq (q-c)/q < \overline{\phi}$ and both mimicking and partial mimicking is non-profitable for all $\phi \in \left[\underline{\phi},(q-c)/q\right]$. Analogously, if $q \geq \left(\overline{\phi}-c\right)/\overline{\phi}$ then $(q-c)/q \geq \overline{\phi}$ and both mimicking and partial mimicking is non-profitable for all $\phi \in \left[\underline{\phi},\overline{\phi}\right]$.

Proof of Lemma 2. Preliminaries: $\overline{\phi}^N(q)$ and $\underline{\phi}^N(q)$ approaches infinity as respective denominator approaches zero. Setting the denominators equal to zero and solving for q gives

$$\overline{q} = \frac{F(E_{01})-F(E_{00})}{1-F(E_{00})} \tag{.3}$$

and

$$\underline{q} = \frac{F(E_{11})-F(E_{10})}{1-F(E_{10})}. \tag{.4}$$

Hence, $\overline{q} \geq \underline{q}$ if and only if 2.10 holds. To see the single crossing property, solving $\overline{\phi}^N(q) = \underline{\phi}^N(q)$ gives the unique solution

$$q^* = \frac{\overline{\phi}(1 - F(E_{01}))(F(E_{11}) - F(E_{10})) - \underline{\phi}(1 - F(E_{11}))(F(E_{01}) - F(E_{00}))}{\overline{\phi}(1 - F(E_{01}))(1 - F(E_{10})) - \underline{\phi}(1 - F(E_{11}))(1 - F(E_{00}))}. \quad (.5)$$

If the denominator in .5 is equal to zero then $\underline{\phi} = \overline{\phi}$, $\underline{q} = \overline{q}$ and $\overline{\phi}^N(q) = \underline{\phi}^N(q)$ for all q.
(i) Suppose both 2.10 and 2.11 holds. Then $\overline{\phi} \geq \underline{\phi}$ and $\overline{q} \geq \underline{q}$. This implies that $\overline{\phi}^N(q) = \underline{\phi}^N(q)$ an even number of times in the interval $[\underline{q}, 1]$. Since q^* is unique whenever one or both of 2.10 and 2.11 holds we conclude that $\overline{\phi}^N(q) \geq \underline{\phi}^N(q)$ for all $q \in [\underline{q}, 1]$. If both 2.10 and 2.11 holds with equality the (i) is trivially true.
(ii) Suppose 2.10 holds strictly and 2.11 doesn't hold. Furthermore, assume $c \geq \underline{c}$. This makes $\overline{\phi} \geq \underline{\phi}$ by Lemma 1(ii) and $\overline{q} > \underline{q}$. It follows that $\overline{\phi}^N(q) = \underline{\phi}^N(q)$ an even number of times in the interval $[\underline{q}, 1]$. q^* is unique in this case and, hence, $\overline{\phi}^N(q) \geq \underline{\phi}^N(q)$ for all $q \in [\underline{q}, 1]$.
(iii) Suppose 2.10 holds strictly and 2.11 doesn't hold. Furthermore, assume $c < \underline{c}$. This makes $\overline{\phi} < \underline{\phi}$ by the proof of Lemma 1 and $\overline{q} > \underline{q}$. It follows that $\overline{\phi}^N(q) = \underline{\phi}^N(q)$ an odd number of times in the interval $[\underline{q}, 1]$. q^* is unique and we conclude that $\overline{\phi}^N(q) \geq \underline{\phi}^N(q)$ for all $q \in [\underline{q}, q^*]$.
(iv) Suppose 2.11 holds strictly, $c \leq \underline{c}$ and 2.10 doesn't hold. Then $\overline{\phi} \geq \underline{\phi}$ by Lemma 1(iii) and $\underline{q} > \overline{q}$. It follows that $\overline{\phi}^N(q) = \underline{\phi}^N(q)$ an odd number times in the interval $[\underline{q}, 1]$. q^* is unique and hence $\overline{\phi}^N(q) \geq \underline{\phi}^N(q)$ for all $q \in [q^*, 1]$.

Proof of Proposition 2. By assumption does (i), (ii), (iii) or (iv) in Lemma 2 hold and $\overline{\phi}^N(q) > \underline{\phi}^N(q)$. Let $\phi \in \left[\underline{\phi}^N(q), \overline{\phi}^N(q)\right]$. Consider L's deviation $s'_L \neq s^S_L$. Suppose s'_L prescribes L to chose an incompetent co-worker, then $s^S_L \succ_{\mathsf{L=0}} s^M_L$ by assumption and $s^M_L \succeq_{\mathsf{L=0}} s'_L$ by the properties of O's belief formation which also gives $s^S_L \succeq s'_L$. To chose an incompetent co-worker is not a part of any strictly profitable deviation. Next, suppose s'_L prescribes a competent co-worker. By assumption is $s^S_L \sim_{\mathsf{L=1}} s^M_L$ and $s^M_L \succeq_{\mathsf{L=1}} s'_L$ by the properties of O's beliefs. Also, $s^S_L \succeq_{\mathsf{L=0}} s'_L$. P's strategy s^S_O maximizes her utility by definition. Given the assumptions above, there exists no strictly profitable deviation and s^S is a NE. ■

Proof of Lemma 3. The observer's expected utility under transparency is

$$u_O\left(s^S\right) = p\left[(1-\theta) + \beta g(1,0)\right] + (1-p)\left[-F(E_{01})\theta + (1 - F(E_{01}))p + \beta g(0,1)\right]$$

and under non-transparency

$$\begin{aligned} u^N_O\left(s^S\right) &= p\left[\left(1-\theta'\right) - (1 - F(E_{10}))(1-q)\left(1 - \theta' - p\right) + \beta g(1,0)\right] \\ &\quad + (1-p)\left[-F(E_{01})\theta' + (1 - F(E_{01}))p + \beta g(0,1)\right]. \end{aligned}$$

Setting $u_O\left(s^S\right) \geq u_O^N\left(s^S\right)$ and simplifying gives

$$\theta' - \theta \geq -\frac{\left(1 - F\left(E_{10}\right)\right)\left(\left(1 - \theta' - p\right) + p\theta'\right)}{p + \left(1 - p\right) F\left(E_{01}\right)}$$

which proves the if statement. The omitted proof of the only-if statement can be reconstructed by doing the "if proof" backwards. Analogously, the competent leader's expected payoff under transparency is

$$u_L\left(s^S \mid \mathsf{L} = 1\right) = \phi - \left(1 - F\left(E_{10}\right)\right) c + g\left(1, 0\right)$$

and under non-transparency

$$u_L^N\left(s^S \mid \mathsf{L} = 1\right) = \left(F\left(E_{10}\right) + q\left(1 - F\left(E_{10}\right)\right)\right)\phi' - \left(1 - F\left(E_{10}\right)\right) c + g\left(1, 0\right).$$

Rewriting $u_L\left(s^S \mid \mathsf{L} = 1\right) \geq u_L^N\left(s^S \mid \mathsf{L} = 1\right)$ gives

$$\phi' - \phi \leq \left(1 - q\right)\left(1 - F\left(E_{10}\right)\right)\phi'.$$

Also here is the proof of the only-if statement omitted.

Proof of Proposition 3. By Lemma 3 it is $\phi' > \phi$ and $\theta' < \theta$ a necessary condition for non-transparency to Pareto dominate transparency. This contradicts the assumption of Weak Monotonicity. ■

References

[1] Besley, T. and Case, A. (1995). “Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits”, *Quarterly Journal of Economics* **110**, 769-798.

[2] Cho, I-K. and Kreps, D.M. (1987). ”Signaling Games and Stable Equilibria”, *Quarterly Journal of Economics* **102**, 179-221.

[3] Effinger, M. and Polborn, M. (1999). “Herding and Anti-Herding: A Model of Reputational Differentation”, mimeo, University of Munich.

[4] Ferejohn, J. (1986). “Incumbent Performance and Electoral Control”, *Public Choice* **50**, 5-26.

[5] Gibbons, R. and Murphy, K.J. (1992). “Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence”, *Journal of Political Economy* **100(3)**, 468-505.

[6] Harrington, J.E. (1993). ”Economic Policy, Economic Performance, and Elections”, *American Economic Review* **83(1)**, 27-42.

[7] Jeon, S. (1996). “Reputational Concerns and Managerial Incentives in Investment Decisions”, *European Economic Review* **42**, 1203-1219.

[8] Levy, G. (1999). “Strategic Consultation and ‘Yes Man’ Advisors” in *Essays on Strategic Information Transmission*, Dissertation in Economics, Princeton University.

[9] Meyer, M.A. and Vickers, J. (1997). “Performance Comparisons and Dynamic Incentives”, *Journal of Political Economy* **105(3)**, 547-581.

[10] Rogoff, K. (1990). ”Equilibrium Political Budget Cycles”, *American Economic Review* **80(1)**, 21-36.

[11] Rogoff, K. and Sibert, A. (1988). ”Elections and Macroeconomic Policy Cycles”, *Review of Economic Studies* **55**, 1-16.

[12] Persson, T. and Tabellini, G. (1990). *Macroeconomic Policy, Credibility and Politics.* London: Harwood Academic Publishers.

[13] Segendorff, B. (1999). ”A Signalling Theory of Scapegoats”, mimeo, Stockholm School of Economics.

[14] Trueman, B. (1994). “Analyst Forecasts and Herding Behavior”, *Review of Financial Studies* **7**, 97-124.