"Yes Men" and "No Men": Does Defiance Signal Talent?

Jeong-Yoo Kim* Keunkwan Ryu

April 16, 2003

Abstract

We provide the rationale for the existence of "yes men" and "no men" in an organization or a group. On one hand, a person is inclined to conform to the instruction of another, because he cannot ignore the information contained in the instruction, even though his own evidence contradicts the instruction. On the other hand, if only the person knows the accuracy of his own information, he may tend to disobey the instruction, to make others believe that he is able in the sense that his information is accurate. We demonstrate that disobedience can signal high ability in an equilibrium. (JEL: D82)

1 *Introduction*

Within an organization, some agents conform to the direction of the supervisor while others do not. In a family, some children obey their parents while others do not. In a high school, there are exemplary students who obey teachers and troublesome students who do not. In a society, some people follow social norms, while others do not hesitate to break them. The former are usually called "yes men" and the latter are called "no men." When the

*We are grateful to the Editor, two anonymous referees, seminar participants at SUNY (Albany), University of Wisconsin (Madison) and audiences at the conference of Korean Economic Association, the Applied Microeconomics Workshop held at the Korea Foundation for Advanced Studies, Far Eastern Econometric Society Meeting held in Kobe for helpful comments. We also wish to acknowledge the financial support from Dongguk University and the nondirected research fund of the Korea Research Foundation.

instruction of another contradicts what he believes to be right, a "yes man" conforms to the opinion of the instructor while suppressing his own opinion, whereas a "no man" often opposes the instructor even when he thinks the instruction is adequate.

Why is there such diversity in behavior? Why are some people more obedient to higher authorities than others? Why do some keep protesting against the superior's decisions even though they may be fired? Social psychologists have argued that the inclination to obey (or to disobey) is determined not only by personal characteristics of individuals, but also by social conditions and organizational environments, such as, gender, age, marital status, educational level, ethnicity, cultural background etc. In this paper, we will address these issues from quite a different perspective, namely from an informational point of view.[1]

Much has been discussed about the rationale for the existence of "yes men." One main rationale is that a person cannot ignore the information others have, because the information contained in the instructor's direction may be useful to himself. This argument corresponds to the explanation of conformity by Bikhchandani, Hirshleifer and Welch [1992] and Banerjee [1992] asserting that a person follows others since he thinks that they know better than himself.

Little, however, has been said about the rationale for "no men," which is the main subject of this paper. We consider an economic agent who cares about the reputation of his ability as well as his monetary return, and whose ability is his private information. In the first best world, a person with more accurate information is more likely to disobey a higher authority, thus, a tendency to say "no" may signal the reliability of a person's information when only the person knows the accuracy of his own information.

To formalize this idea, we consider an organization in which all the decisions are made in a decentralized way. In our framework, the superior makes only an unbinding suggestion to his subordinate. Taking this into account, the subordinate makes a real decision. Therefore, the superior's suggestion is simply a means of conveying information,[2] as long as it is unbinding.

[1]Many empirical studies have found the relation between personality and obedient behavior to be weak or nonexistent. (Marlowe and Gergen [1970]; Mischel [1968]) These findings may confer relative gravity on this work attempting to explain conforming behavior in the informational context.

[2]A payoff-irrelevant message is usually called "cheap talk." It is well-known that cheap talk, even though it is payoff-irrelevant, can reveal some useful information unless the

In this paper, the superior and the subordinate receive independent noisy signals, conditional on the state. Since noises are independent, the signals they receive are, in general, different and this difference is the main cause for disagreement between them. The subordinate is either more or less able than his superior in terms of the information's accuracy, and he derives utility directly from being perceived as able. Differences in the ability of subordinates, then, lead to differences in their tendency to be disobedient. Whether or not a subordinate brings the disagreement into effect relies totally on this difference in the accuracy of signals. A subordinate whose information is more accurate tends to be more disobedient because he puts more weight on his own information when it appears to conflict with the superior's information. It is also the difference in the accuracy of information that enables an able subordinate to be separated from a less able subordinate who wants to pretend to be able by being disobedient. Separation is possible because a difference in ability differentiates the expected monetary loss from disobedience.

We characterize pooling equilibria and a separating equilibrium. In a pooling equilibrium, a low-ability subordinate succeeds in mimicking a high-ability subordinate. Intuitively, a pooling equilibrium is possible when a person's deviation from the socially accepted standard of disobedience gives the impression that he is less able, and the resulting reputational loss exceeds the expected monetary gain from the deviation regardless of his true ability. In this equilibrium, a subordinate may be a "no man" for fear that obedience may be interpreted as "less able." Also, we find a separating equilibrium in which the able subordinate is more inclined to disobey than the less able one.[3] In this equilibrium, a less able subordinate cannot mimic an able subordinate perfectly because of a difference in the signalling cost. Since the information of a less able subordinate is noisier, he cannot rely on his own information enough to disobey as frequently as an able subordinate. After all, if the subordinate is motivated by reputational concerns, some subordinate behaves too disobediently overriding his own information then, as a result, some inefficiency occurs both in a pooling equilibrium and in a

preferences of the sender and the receiver are quite opposite. (Crawford and Sobel [1982])

[3]There are some works in social psychology suggesting that people who have high self-esteem may be particularly disobedient. (Asch [1956]; Crutchfield [1955]) Also, it has been reported that a person is less likely to obey when the instructor is challenged by his authority, expertise, competency or judgment etc. (Croner and Willis [1961])

separating equilibrium. This may tell us, for example, why some agents in a brokerage firm like Merrill Lynch give their clients financial advice against recommendations from the top research department.

In the meantime, the subordinate's decision depends crucially on the superior's strategic motive, because the superior can transmit his information strategically in a way that is favorable to him. For example, if the superior believes that the subordinate will be very likely to act against his suggestion, he will find it in his interest to make a reverse suggestion, the opposite of what he wants done. Then, the subordinate will adjust his inference and its consequent decision accordingly. However, this possibility does not generate a serious problem in the sense that the superior's true information is transmitted in equilibrium, whether his suggestion is reverse or not.

Although we employ an organizational framework throughout, the notions of "yes men" and "no men" in general do not require a hierarchical structure. The explanation is also applicable to those who are more (or less) likely to comply with their friends or colleagues.

In deriving the results, the reputation-formation process plays a crucial role. The subordinate's neighbors, for example, his colleagues cannot observe the final output consequent on the subordinate's choice. They can only observe his tendency to be disobedient and base their judgment of his ability on this observation. It is assumed that a subordinate is not concerned with his superior's perception of his ability. A higher authority's perception may matter to a person in an organization because it may affect his salary. But in a non-organizational community, the perception of colleagues is most crucial because there is more interaction with them.

The paper is organized as follows. We review the related literature in Section 2. Next, we introduce a simple model to illustrate the intuition in Section 3. In Section 4, we present a general model. In Section 5, as a benchmark case, we analyze the model when the subordinate's ability is known. In Section 6, we characterize equilibria under incomplete information. Discussion and concluding remarks follow in Section 7.

2 Related Literature

There is some literature on reputational cheap talk which is closely related to this paper.

Scharfstein and Stein [1990], which is seminal in this literature, considers two managers who make investment decisions sequentially. A manager's decision affects his payoff only through the market's belief about his ability. In this sense, an investment decision is cheap talk. In their model, smart managers receive correlated informative signals about the value of an investment, while dumb managers receive independent, purely noisy signals. This correlation of signals drives managers to herd in order to share the blame.

Trueman [1994] was the first to establish the anti-herding result similar to the result herein, by using a model of reputational cheap talk. He also considers sequential decisions of two security analysts. In his model, an (intellectually) weak analyst who receives an extreme signal sometimes deflates his forecast, while a strong analyst announces in correspondence with his own information. This is because a weak analyst who cares about his reputation is suspicious that the extreme signal may be in error. As a result, a weak analyst is differentiated from a strong analyst. Ottaviani and Sørensen [2001] obtained a similar result in a committee setting by focusing on the most informative equilibrium. Both models also predict that, in a sequential setting, the tendency of a weak second mover to distort his message from his own judgment becomes stronger because the message from the first mover can provide additional useful information to the second mover. However, in both models, a strong type does not deviate from his first best behavior to separate himself from a low type, which is usual in a cheap talk model in which different messages involve no difference in direct signalling costs.

The most important departure of the model herein from the literature on reputational cheap talk is that one who assesses a person's ability cannot observe the realization of the final output at least before he makes an assessment of the ability. So, he cannot use the information of the realized output to update his belief. This may make separation between a strong type and a weak type difficult. However, in our model, a subordinate's action is not cheap talk, because he does receive a share of the actual output. This facilitates separation. Moreover, as a result, a strong subordinate can separate himself more effectively by distorting his action in a way to make it difficult for a weak subordinate to mimic him. Consequently, even a strong subordinate's behavior is affected by his predecessor in our model, contrary to Trueman [1994] and Ottaviani and Sørensen [2001]. In fact, differentiation in those models are not between a weak type and a strong type, in the sense that such difference can occur even when two players move simultaneously. The main interest of this paper is in differentiated behavior between sequential movers.

LEVY [2000] provides a model in which a decision maker receiving a private signal about the state of the world considers whether to consult an adviser in order to gather more information. Her model is similar to ours in that an adviser can misrepresent his information strategically. Her main intuition is that the conflict between career concerns of the decision maker and the adviser makes an able decision-maker either not to consult an adviser or to ignore his advice. This is not the case in the current model, because, depending on his information, a high-ability subordinate may follow the superior's recommendation. In this model with a richer signal space, the superior's decision to be truthful is made by comparing two probabilities that his recommendation will be followed, and that it will be disobeyed, based on the expectation of the subordinate's information.

There is some related literature. PRENDERGAST [1993] considers the possible incentive for a subordinate to become a "yes man" in a quite different context. In his model, a subordinate whose role is to collect information conforms to the opinion of the superior, since firms generally resort to subjective evaluation procedures, where a superior compares his own finding to that of his subordinate.

PRENDERGAST AND STOLE [1996] seem to be the closest to our model in spirit. They consider a manager who makes investment decisions over time. A manager receives a noisy signal each period about the return of investments before he makes a decision and updates his posterior about the return based on this signal. If the manager is able, in the sense that the information he receives has high precision, the posterior of the return is more subject to variation initially, because he puts more weight on the information he has recently received, and gets less subject to variation over time, because his accumulated observations are more reliable than the current observation. Therefore, if his ability is his private information and he cares about his reputation for high ability, the manager will initially overreact to new information but, ultimately, become too conservative to change his investments. Thus, varying investments at an early stage signals talent in their model. As in our model, they assume that a manager is concerned about profits as well as his reputation, so that his decision is not cheap talk. Without preferences over profits, no separating would occur in their model, which is the case in our model. Despite many common features, our paper provides a new insight by considering another strategic economic agent who is in a position to control the information available to a person making a payoff-relevant decision.

3 *Illustrative Model*

3.1 Setup

Suppose there is one superior or chief (c) and many subordinates including one representative subordinate (s) in an organization. A superior recommends his subordinate either to do ($a = 1$) or not to do ($a = 0$). If the subordinate chooses to do, it yields a random profit θ_1, which takes on either H (good state) or L (bad state). The prior probabilities of H and L are μ and $1 - \mu$, respectively. Not doing ($a = 0$) yields the sure return $\theta_0 = 0$. After receiving the superior's recommendation, the subordinate decides whether to follow the recommendation. The decision by the subordinate is payoff-relevant. We assume that

$$\mu H + (1 - \mu)L = 0, \tag{1}$$

which means that the subordinate is *a priori* indifferent between doing and not doing.

Before making choices, each of the superior and the subordinate receives a private signal, v_i, $i = c, s$, which can take on one of two values: H (good signal) or L (bad signal). We assume that both of v_c and v_s are informative. That is, a good (bad) signal is more likely to occur in a good (bad) state than in a bad (good resp.) state. Formally, we assume:

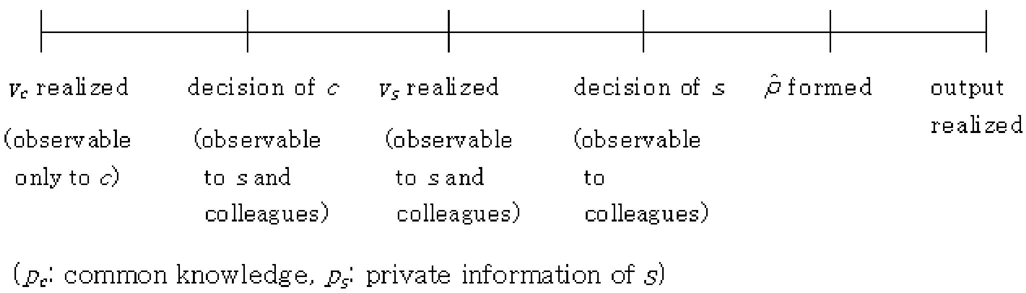
$$\text{Prob}(v_i \mid \theta_1) = \text{Prob}(H \mid H) = \text{Prob}(L \mid L) = p_i > 1/2, i = c, s. \tag{2}$$

p_i can be interpreted as the informativeness of the signal player i receives. We assume that p_c is known to both c and s, while p_s is private information of s. p_s is assumed to be either p_H or p_L with $p_L < p_c < p_H$. If $p_s > p_L$ (i.e. $p_s = p_H$), which occurs with prior probability ρ, we can say that the subordinate has more accurate information than the superior has, or put differently, that the subordinate knows better than the superior. If $p_s < p_c$, the subordinate can be said to have a poorer knowledge than his superior.[4] Henceforth, we will call a subordinate with p_H a high-ability type (H) and a subordinate with p_L a low-ability type (L).

[4] Knowledge specific in a certain job can be thought of as acquired ability, although it is not inherent ability. Better knowledge usually enables a person to process information he receives in a more refined way and to extract a better signal underlying the information.

The colleagues update their posterior beliefs about the subordinate's ability, based on their observations. We assume that the colleagues cannot observe the output at least before they form the posterior beliefs. It usually takes considerable time to observe the final output. Also, since v_c and v_s are private information of the superior and the subordinate respectively, they are not observable, nor communicable to colleagues. So, all that the colleagues can observe and base their belief-updating on is the subordinate's choice of a. We denote by $\hat{\rho}$ the posterior belief that $p_s = p_H$. The sequence of events is illustrated in Figure 1.

Figure 1
Sequence of Events and Information Structure



We assume that the superior and the subordinate divide the final outcome by a predetermined sharing ratio, $1-\alpha$, α, and that the subordinate gets not only direct utility from his income, but also some utility from being thought of as an expert by his colleagues. The utility function of the superior is assumed to be $U^c(a) = (1-\alpha)\theta_a$. Similarly, the utility function of the subordinate is assumed to be $U^s(a) = \alpha\theta_a + k\hat{\rho}$ where $k(>0)$ is the weight that the subordinate attaches to his reputation, relative to the direct utility from his income. Here, $(1-\alpha)\theta_a$ and $\alpha\theta_a$ are intrinsic utility of c and s respectively, and $k\hat{\rho}$ is the extrinsic utility of s.

3.2 Bayesian Updating

The subordinate can infer some valuable information about θ_1 from the superior's direction. Then, his assessment can be made based on this inferred information as well as his own information, v_s. Thus, Bayes' law allows the subordinate to modify the probabilities attached to the good state: If the superior's direction implies that $v_c = H$, the probabilities are

$$\text{Prob}(\theta_1 = H \mid v_c = H, v_s = H) = \frac{\mu p_c p_s}{\mu p_c p_s + (1-\mu)(1-p_c)(1-p_s)} \equiv \beta_{HH}(p_s); \tag{3}$$

$$\text{Prob}(\theta_1 = H \mid v_c = H, v_s = L) = \frac{\mu p_c (1-p_s)}{\mu p_c (1-p_s) + (1-\mu)(1-p_c)p_s} \equiv \beta_{HL}(p_s). \tag{4}$$

If the superior's direction conveys the informataion that $v_c = L$, the probabilities are

$$\text{Prob}(\theta_1 = H \mid v_c = L, v_s = H) = \frac{\mu (1-p_c) p_s}{\mu (1-p_c) p_s + (1-\mu)p_c(1-p_s)} \equiv \beta_{LH}(p_s); \tag{5}$$

$$\text{Prob}(\theta_1 = H \mid v_c = L, v_s = L) = \frac{\mu (1-p_c)(1-p_s)}{\mu (1-p_c)(1-p_s) + (1-\mu)p_c p_s} \equiv \beta_{LL}(p_s). \tag{6}$$

Observe that $\beta_{HH}(p_s)$ and $\beta_{LH}(p_s)$ are increasing in p_s, whereas $\beta_{HL}(p_s)$ and $\beta_{LL}(p_s)$ are decreasing in p_s. Since $p_c, p_s > 1/2$, it follows that

$$\beta_{LL}(p_s) < \beta_{HL}(p_s), \beta_{LH}(p_s) < \beta_{HH}(p_s), \text{ for all } p_s. \tag{7}$$

Also, it follows that

$$\beta_{LH}(p_L), \beta_{HL}(p_H) < \mu < \beta_{LH}(p_H), \beta_{HL}(p_L) \tag{8}$$

from the assumption that $p_L < p_c < p_H$.

3.3 Full Information

The subordinate's behavior under full information about p_s will provide a good guideline for his behavior when p_s is his private information.

Due to the sequential nature of decisions, the subordinate can infer the superior's signal from his recommendation.

Suppose the superior recommends $a = 1$ if $v_c = H$ and $a = 0$ if $v_c = L$. In this benchmark case of full information, any action taken by the subordinate cannot affect $\hat{\rho}$, so that he will simply choose an action to maximize the expected profit, based on his information and the superior's recommendation.

If the superior recommends $a = 0$, the subordinate's decision is determined by the sign of $E[\theta_1 \mid v_c = L, v_s; p_s]$. From (1) and (8), one can show that,

$$E[\theta_1 \mid v_c = L, H; p_H] = \beta_{LH}(p_H)H + [1 - \beta_{LH}(p_H)]L > 0, \tag{9}$$

$$E[\theta_1 \mid v_c = L, H; p_L] = \beta_{LH}(p_L)H + [1 - \beta_{LH}(p_L)]L < 0. \tag{10}$$

This implies that, under full information, L-type subordinate follows the superior's recommendation and H-type subordinate does not, when he has received a good signal.

It is clear from (7) that both types of subordinate having received a bad signal follow the superior's recommendation.

The analysis for the case that $a = 1$ is recommended is similar, so it will be omitted.

3.4 Asymmetric Information

The analysis above suggests the possibility that, if the subordinate's ability is not known to others, one who tends to disobey the superior's instruction may be thought as an expert, an able man. So, a subordinate who is eager to get a reputation for high ability may strategically disobey the superior's instruction even though his statistical inference favors obeying it.

The solution concept that will be employed is the weak Perfect Bayesian equilibrium (WPBE). To illustrate, we focus on the continuation game given that the superior has made its decision of $a = 0$.

In order to make the analysis interesting, we assume that the first-best outcome achieved under full information cannot be supported as an equilibrium under asymmetric information. This is equivalent to assuming that

$$\alpha\beta_{LH}(p_L)H + \alpha[1 - \beta_{LH}(p_L)]L + k > 0. \tag{11}$$

This implies that L-type subordinate can successfully mimic H type if H-type subordinate behaves as under full information. Then, we have

Observation 1 In the continuation game, there is a semi-separating equilibrium in which H-type subordinate with a signal H does not follow the superior and L-type subordinate randomizes with some probability when he observes H, while both types with a signal L follow the superior.

In order to prove this, it is sufficient to find a randomizing probability which makes L-type subordinate indifferent between choosing 0 and 1 when he observed H. Let r be the probability that he chooses $a = 1$. Denoting posterior beliefs when he chooses $a = 0, 1$ by $\hat{\rho}_0$, $\hat{\rho}_1$ resp., r must satisfy

$$\alpha\beta_{LH}(p_L)H + \alpha[1 - \beta_{LH}(p_L)]L + k\hat{\rho}_1 = k\hat{\rho}_0. \tag{12}$$

Here, $\hat{\rho}_0$ and $\hat{\rho}_1$ can be computed as

$$\hat{\rho}_0 = \frac{\rho(1-\mu)}{(1-\rho)(1-\mu r) + \rho(1-\mu)}, \tag{13}$$

$$\hat{\rho}_1 = \frac{\rho}{\rho + (1-\rho)r}, \tag{14}$$

by using the Bayes' law.

Let $\tilde{\rho}(r) = \hat{\rho}_1(r) - \hat{\rho}_0(r)$. Then, $\tilde{\rho}(1) = 0$ and $\tilde{\rho}(0) = (1-\rho)/(1-\rho\mu)$. Thus, by continuity, there exists $r^* \in (0, 1)$ satisfying (12) if $k > \rho_\mu B_{LH}$, where $B_{LH} \equiv -\alpha[\beta_{LH}(p_L)H + \{1 - \beta_{LH}(p_L)\}L] > 0$ and $\rho_\mu = (1-\rho\mu)/(1-\rho)$.

On the other hand, in order for a subordinate who observes L not to deviate, it must be the case that:

$$\alpha\beta_{LL}(p_s)H + \alpha[1 - \beta_{LL}(p_s)]L + k\hat{\rho}_1 \leq k\hat{\rho}_0, \text{ for } p_s = p_L, p_H. \tag{15}$$

Since it is more attractive for L-type subordinate to deviate by disobeying after observing L, (15) is reduced to

$$k \leq \rho_\mu B_{LL}, \tag{16}$$

where $B_{LL} \equiv -\alpha[\beta_{LL}(p_L)H - \{1 - \beta_{LL}(p_L)\}L] > 0$. Since $B_{LH} < B_{LL}$, there is a proper interval of k, $[\rho_\mu B_{LH}, \rho_\mu B_{LL}]$, which makes the above semi-separating equlibrium viable. This suggests that disobedience can signal his ability.

We have not yet demonstrated that this signaling outcome constitutes part of an equilibrium in the overall game including the superior's decision, because this section is designed only for an illustrative purpose. We will perform this task in a more general model.

4 General Model

In this section, we generalize our argument by considering a continuum of possible outcome and a continuum of signal values instead of the binary outcome and binary signals.

Suppose that θ_a is normally distributed with mean μ_a and precision $h_a(> 0)$ (i.e., variance $1/h_a$). We will assume that θ_0 and θ_1 have the same mean and normalize it to 0, $\mu_0 = \mu_1 = 0$. Besides, each of the superior and the subordinate receives his noisy information v_i on θ_1 right before making a decision. We will assume the following.

$$v_i = \theta_1 + \epsilon_i, \tag{17}$$

where ϵ_i is normally distributed with zero mean and precision $h_i(> 0)$, $i = c, s$. θ_1, ϵ_c and ϵ_s are mutually independent. h_c is assumed to be known to both c and s, whereas h_s is assumed to be private information of s. Also, we assume that h_s is either h_H or h_L with $h_L < h_c < h_H$ and that the prior probability that $h_s = h_H$, denoted by $\rho \in (0, 1)$, is common knowledge.

As in Section 3, we assume that v_c is private information of the superior and is not directly communicated to the subordinate.[5] But, unlike in Section 3, we assume that v_s is known to the colleagues as well as the subordinate. Usually, colleagues around the subordinate are in a good position to observe the information he collects. So, it is common that the colleagues know on what ground a subordinate is protesting against the superior as well as the simple fact that he is protesting.

For our analysis, we will make formal definitions of a "yes man" and a "no man" below.

Definition 1 A subordinate is said to be a "yes man" if he conforms to the instruction of his superior even if his own information shows evidence against it. Likewise, a subordinate is said to be a "no man" if he objects to the instruction of his superior even if his own information shows evidence for it.

According to this definition, we do not say that a subordinate is a "no man" simply because he does not follow the instruction of his superior when his own information shows evidence against it.

[5] Huge transaction costs are a typical reason for no direct communication between the superior and the subordinate. In fact, however, this is not an essential assumption. Even though v_c is perfectly communicated to the subordinate, the qualitative nature of the result will be unaffected.

5 Full Information and a Rationale for "Yes Men"

As asserted in Section 3, the superior's recommendation conveys some information which is also useful to the subordinate. This is the reason why a subordinate sometimes obeys the superior against his own information.

For the time being, suppose the superior recommends $a = 0$ if and only if $v_c \leq 0$. Then, the subordinate who is recommended to choose $a = 0$ will conform to it, if

$$E[\theta_1 \mid v_c \leq 0, v_s; h_s] \leq 0. \tag{18}$$

Since $E[\theta_1 \mid v_c \leq 0, v_s]$ is increasing in v_s,[6] the subordinate obeys if $v_s \leq v^0$ and disobeys if $v_s > v^0$, where $E[\theta_1 \mid v_c \leq 0, v^0; h_s] = 0$. This is just because a higher value of v_s implies a higher expected value of θ_1 regardless of h_s. Notice that v^0 is a function of h_c and h_s. The following propositions will be useful for future analysis.

Proposition 1 $v^0(h_c, h_s) > 0$ *for all* h_c, h_s.

Proof. See the appendix.

This proposition says that the subordinate with $v_s = 0$ will strictly prefer to conform to the superior. This implies that a subordinate may behave as a "yes man." The intuition is crystal clear. Since the superior's instruction of $a = 0$ implies that he has a negative signal, $v_c \leq 0$, a positive signal is required to offset the effect of this information and to make $a = 1$ profitable.

Proposition 2 (i) $\partial v^0(h_c, h_s)/\partial h_s < 0$. *(ii)* $\partial v^0(h_c, h_s)/\partial h_c > 0$.

Proof. straightforward

This proposition implies that a subordinate is less inclined to conform to the superior as his own information gets more precise, or as the information of the superior is less precise.

The analysis for the case that $a = 1$ will be omitted, since it is qualitatively the same.

[6] For a rigorous proof, see Lemma 1.

For the analysis of the model under asymmetric information about h_s, the following lemma will turn out to be useful.

Lemma 1 (i) $E[\theta_1 \mid v_c \leq 0, v_s, h_s]$ *is increasing in* v_s. *(ii)* $\partial E[\theta_1 \mid v_c \leq 0, v_s, h_s]/\partial v_s$ *is increasing in* h_s. *(iii)* $E[\theta_1 \mid v_c \leq 0, v_s = 0, h_L] < E[\theta_1 \mid v_c \leq 0, v_s = 0, h_H] < 0$. *(iv)* $\lim_{v_s \to -\infty} E[\theta_1 \mid v_c \leq 0, v_s, h_L] > \lim_{v_s \to -\infty} E[\theta_1 \mid v_c \leq 0, v_s, h_H]$.

Proof. See the appendix.

It is intuitively clear that the expected value of θ_1 for type H has a higher slope with respect to v_s. For type H puts more weight on the given information v_s in calculating the expected output when he chooses a_1. A good news ($v_s > 0$) is a better signal to type H, while a bad news ($v_s < 0$) is a worse signal to him. This observation makes it also clear that, if v_s is very small, type L expects a higher output when he chooses $a = 1$ than type H, whereas type H expects a higher output if v_s is very large. This implies that there must be $\tilde{v_s}$ such that $E[\theta_1 \mid v_c \leq 0, \tilde{v_s}, h_H] = E[\theta_1 \mid v_c \leq 0, \tilde{v_s}, h_L] \equiv g(< 0)$. Notice that $\tilde{v_s} < 0$ from (iii) and (iv) of Lemma 1 and continuity. Henceforth, for our analysis, we will make the following assumption.

Assumption 1 $k + \alpha g \leq 0$.

This assumption simply says that the reputational effect (k) is not so large relative to the real output effect ($-\alpha g$).

6.1 Subordinate's Decision

Without reputatioinal considerations, i.e., $k = 0$, the high-ability subordinate would conform to the recommendation of $a = 0$ if $v_s \leq v_H$ and the low-ability subordinate would conform if $v_s \leq v_L$ where v_L and v_H are defined as satisfying (18) with equality in cases that $h_s = h_L$, h_H respectively. Notice that $v_H < v_L$ by Proposition 2. In other words, the high-ability subordinate is more likely to disobey. Thus, if a subordinate with $v_H \leq v_s < v_L$ disobeys, his colleagues will believe that the subordinate is of high ability with probability one ($\hat{\rho} = 1$), and if he conforms, they will believe that he is of low ability with probability one ($\hat{\rho} = 0$). If $v_s > v_L$ or $v_s \leq v_H$, nothing can be inferred about the subordinate's ability from his decision and thus the prior belief is preserved i.e., $\hat{\rho} = \rho$.

With reputational considerations, however, decision rules are altered. There are two possible kinds of equilibria in this game; separating equilibria in which each type of the subordinate adopts a different strategy rule i.e., chooses a different action for some v_s,[7] and pooling equilibria in which each type adopts the same strategy rule, i.e., chooses the same action for all v_s.

For the moment, we will restrict our attention to strategies with the cutoff-rule property that each type of the subordinate does not obey the superior if and only if his noisy signal v_s exceeds a threshold level.

Let $\hat{v_L}$ and $\hat{v_H}$ be the cutoff values of type L and H. Then, formally, $(\hat{v_L}, \hat{v_H})$ constitutes an equilibrium if s-type subordinate disobeys if $v_s > \hat{v}_s$ and obeys if $v_s \leq \hat{v}_s$, for all $s = H, L$. If $\hat{v_L} \neq \hat{v_H}$, it is a separating equilibrium, whereas it is a pooling equilibrium, if $\hat{v_L} = \hat{v_H}$.

Separating Equilibrium Our main concern is whether there is a separating equilibrium in which one can infer a subordinate's ability from his inclination towards disobedience. Theorem 1 shows, however, that there is no separating equilibrium in this game.

Theorem 1 There is no separating equilibrium.

Proof. See the appendix.

The intuitive reason is as follows. In a separating equilibrium, the loss in the intrinsic utility from choosing $a = 1$ must exceed the reputational gain (k) to the low-ability subordinate regardless of his information within the separating interval $(\hat{v_H}, \hat{v_L}]$, while the latter ($k\rho$) must exceed the former for any $v_s \in (\hat{v_L}, \infty)$. Since the loss in the intrinsic utility varies continuously with v_s, it is not possible so long as there is a discrepancy between k and $k\rho$.[8]

Pooling Equilibrium We next turn to pooling equilibria.

[7]Unlike in standard signaling games, in our model, the type of the subordinate may not be revealed in a separating equilibrium for some v_s.

[8]More optimistic beliefs can be assigned off the equilibrium path. However, for the purpose of demonstrating the nonexistence of the separating equilibrium, it suffices to consider the most pessimistic one.

Let the pooling equilibrium cutoff value be $\hat{v}$, so that a subordinate with $v_s > \hat{v}$ chooses $a = 1$ regardless of his ability, whereas a subordinate with $v_s \leq \hat{v}$ chooses $a = 0$ regardless of his ability in equilibrium. To characterize the set of equilibrium values for $\hat{v}$, we need to find the posterior beliefs to support them as equilibrium outcomes. The colleagues must preserve their prior beliefs, after observing a subordinate with $v_s > \hat{v}$ disobeying or observing a subordinate with $v_s \leq \hat{v}$ obeying. Also, we impose the most pessimistic belief of $\hat{\rho} = 0$ off the equilibrium path. This pessimistic belief will help us to characterize the whole set of pooling equilibria.

From incentive compatibility conditions, $\hat{v}$ must satisfy

$$\alpha E[\theta_1 \mid v_c \leq 0, v_s, h_L] + k\rho \geq 0, \forall v_s > \hat{v}, \tag{19}$$

$$\alpha E[\theta_1 \mid v_c \leq 0, v_s, h_H] \leq k\rho, \forall v_s \leq \hat{v}.[9] \tag{20}$$

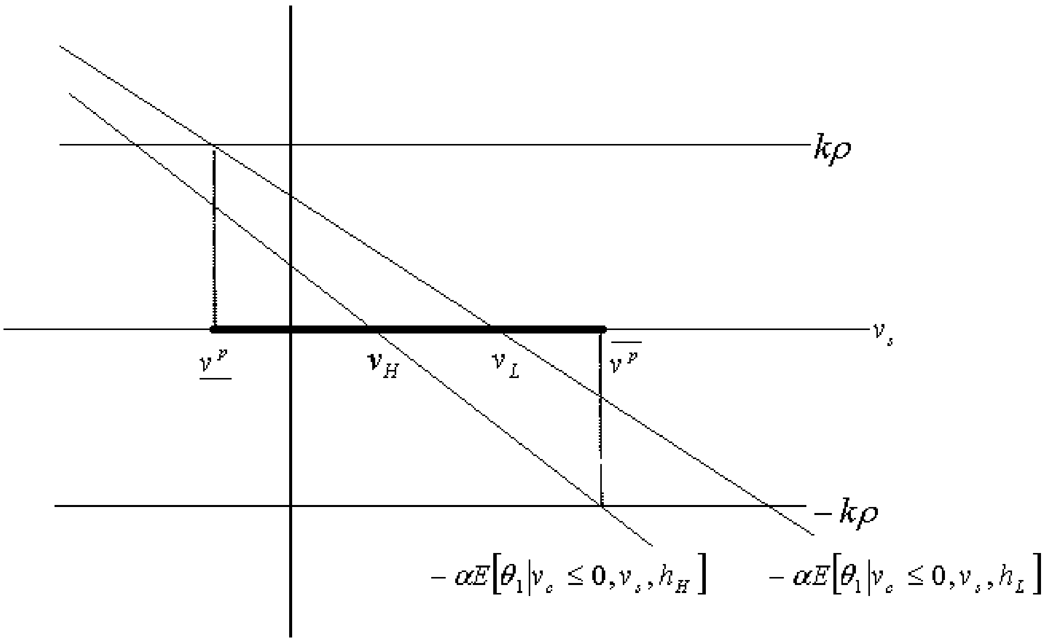
Inequality (19) requires that a low type of subordinate with $v_s > \hat{v}$ prefers disobeying the instruction, and inequality (20) requires that a high type of subordinate with $v_s \leq \hat{v}$ is willing to obey the superior. Let $\underline{v^p}$, $\overline{v^p}$ be v_s satisfying (19) and (20) with equality. Then, any $v_s \in [\underline{v^p}, \overline{v^p}]$ can be a pooling equilibrium cutoff value. As long as $\hat{v} \geq \underline{v^p}$, the reputational gain from disobeying exceeds the loss in the intrinsic utility from it to both types of subordinate with $v_s > \hat{v}$, and, as long as $\hat{v} \leq \overline{v^p}$, the reputational gain from obeying is greater than the loss in the intrinsic utility from obeying (the gain from disobeying) to both types of subordinate with $v_s < \hat{v}$. The set of all the pooling equilibrium cutoff values $\hat{v}$ are drawn as a bold line in Figure 2.

Theorem 2 There exists a continuum of pooling equilibria in which a subordinate makes the same choice regardless of his ability for any information they receive, if there is a nonempty interval $[\underline{v^p}, \overline{v^p}]$. In a pooling equilibrium, a subordinate may be a "no man" in the sense that he may not follow the instruction of $a = 0$ even if $v_s < 0$.

Proof. See the appendix.

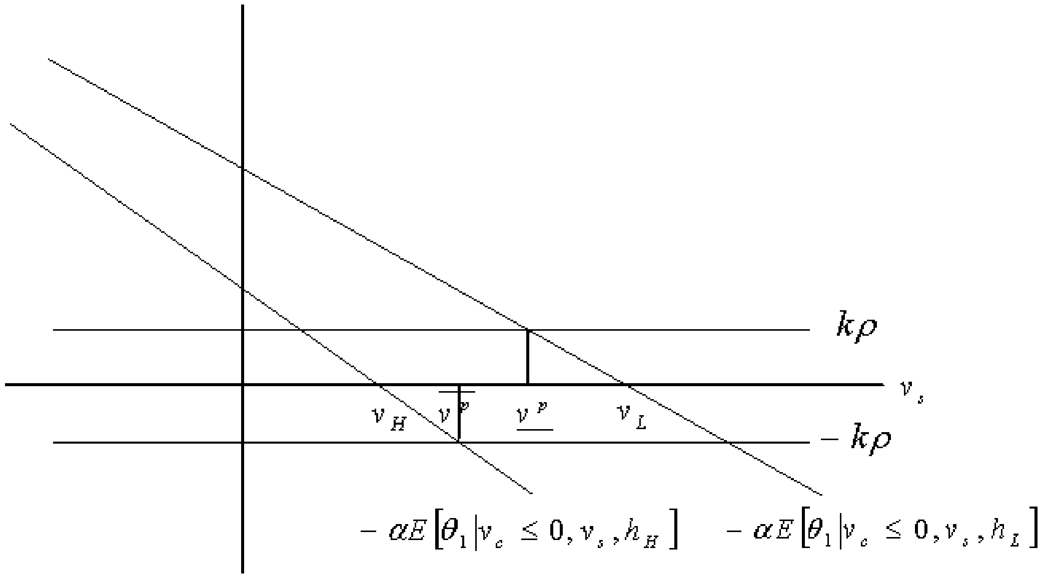
[9]The other two incentive compatibility conditions are redundant.

Figure 2
Set of Pooling Equilibria



In a pooling equilibrium, both types behave in the same way, so that disobedient behavior cannot be considered as a signal of high ability. In this equilibrium, a subordinate is sometimes disobedient even if his own information, together with the superior's information inferred, favors obeying because, otherwise, he would be regarded as less able. From Figure 2, one can see that, as the subordinate cares more about his reputation, i.e., k becomes larger, the set of pooling equilibria is enlarged, and, as a result, a subordinate is more likely to be a "no man." On the contrary, if the subordinate cares little about his reputation, the set of pooling equilibria may be very small, or, in the extreme, pooling equilibria may not exist at all. (See Figure 3.) This is because a deviation from the equilibrium action would not entail a great punishment. Similarly, if the proportion of high-ability subordinates in an organization is higher (ρ is larger), the set of pooling equilibria becomes larger. This is also because not conforming to the equilibrium action would damage the subordinate's reputation more severely.

Figure 3
Possibility of No Pooling Equilibrium



As a final note, not all the pooling equilibria are reasonable. To eliminate unreasonable outcomes, we will appeal to a more refined concept, so-called the universally divine equilibrium by Banks and Sobel [1987]. Roughly, universal divinity requires off-the-equilibrium beliefs to place positive probability only on those types that are most likely to deviate. Here, we say that a type H is more likely to deviate to saying "no", since whenever any given belief would lead a type L to deviate to saying "no", a type H would as well, but not vice versa. Thus, an equilibrium value, $\hat{v}$, satisfies universal divinity if, for all $v_s (\leq \hat{v})$,

$$\alpha E[\theta_1 \mid v_c \leq 0, v_s, h_s] + k \leq k\rho, \ s = L, H, \tag{21}$$

i.e., neither type of subordinate with $v_s \leq \hat{v}$ has an incentive to disobey when disobeying makes him perceived as H type. For our purpose, let us define v_U as v_s satisfying $\alpha E[\theta_1 \mid v_c \leq 0, v_U, h_H] + k(1-\rho) = 0$. In other words, v_U is the value of v_s making H-type's expected gain from disobeying zero under this optimistic belief. Then, for any $v_s \leq v_U$, inequality (21) is satisfied. Consequently, both types will prefer obeying for any $v_s \leq \hat{v}$ as long as $\hat{v} \leq v_U$, since their expected gains from disobeying are negative. Also, notice that $v_U < v_H$, as long as $\rho < 1$. Intuitively, since the defying subordinate earns some reputational gain under universal divinity, the expected gain from disobeying is made zero for a smaller value of v_s. So, we have the following theorem.

Theorem 3 A pooling equilibrium cutoff value $\hat{v}(\leq v_U)$ satisfies universal divinity, where $v_U < v_H$.[10] *In a universally divine pooling equilibrium, both types are necessarily more likely to disobey the instruction than under full information.*

Proof. direct from the definitions of v_U and universal divinity

From the analysis above, we can see that inefficiency occurs if the subordinate is of type L (H resp.) and $v_s \in (v_U, v_L)$ (or $v_s \in (v_U, v_H)$ resp.). This inefficiency comes from the colleagues' lack of information about his ability. If the subordinate did not have to care about his reputation, he would behave as under full information about his ability, which would be the first best outcome.

Separating Equilibrium without the Cutoff Rule Property Although we demonstrated that there existed no separating equilibrium by artificially restricting our attention to strategies with the cutoff-rule property, we can obtain a separating equilibrium, if we lift the restriction.

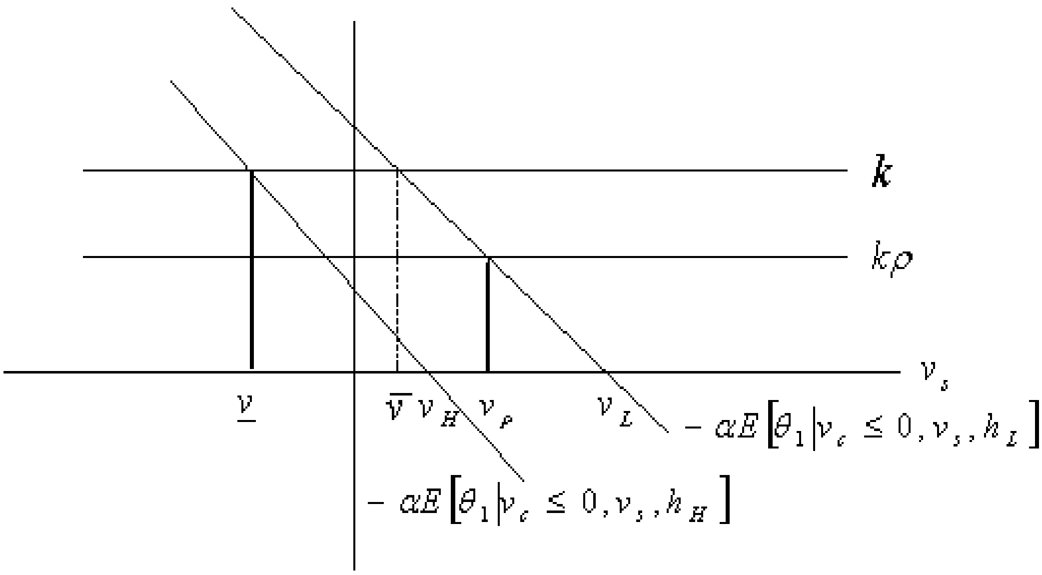
For our purpose, we will allow behavioral strategies. Let $r_s(v_s)$ be the probability that the s-type subordinate with v_s disobeys the instruction of $a = 0$. Suppose the high-ability subordinate adopts the (behavioral) strategy $r_H(v_s) = 0$ if $v_s \leq \hat{v_H}$, $r_H(v_s) = 1$ if $v_s > \hat{v_H}$, and the low-ability subordinate uses the strategy $r_L(v_s) = 0$ if $v_s \leq \hat{v_L}$, $r_L(v_s) = r(v_s)$ for some $r(v_s) \in [0, 1]$ if $\hat{v_L} < v_s \leq v_\rho$ and $r_L(v_s) = 1$ if $v_s > v_\rho$ where $\alpha E[\theta_1 \mid v_c \leq 0, v_\rho, h_L] + k\rho = 0$. (See Figure 4.) The following theorem demonstrates that there is such an equilibrium.

Theorem 4 There exists a separating equilibrium in which, for some values of v_s, the low type subordinate disobeys the superior with probabilities (increasing with v_s but) less than one, while the high type disobeys him with probability one. Furthermore, it always exists if Assumption 1 is valid.

[10] If $\rho < 1/2$, no pooling equilibrium satisfies universal divinity, since there is no $v_U \in [\underline{v}^p, \overline{v}^p]$.

Proof. See the appendix.

Figure 4
Separating Equilibrium without the Cutoff Property



In this equilibrium, defiance can signal high ability. The intuition behind Theorem 4 is as follows. It is easier for a low-ability subordinate with a higher value of v_s to mimic a high-ability one, since the loss from disobeying is smaller. So, to frustrate mimicry, colleagues must believe the disobedient to be of low type with a higher probability if he has a higher value of v_s. Then, for this belief to be consistent, a low-ability one must disobey more frequently as v_s is higher.

Also, notice that, in this equilibrium, a high-ability subordinate distorts his behavior from the first-best outcome by being more disobedient. (See Figure 4.) As in most literature on signaling, the high-ability subordinate and the low-ability one have different signaling costs in our model. The difference in signaling costs is measured by the difference in $-\alpha E[\theta_1 \mid v_c \leq 0, v_s, h_s]$, $s = L, H$. This difference makes it possible for type H to separate himself from type L successfully.

6.2 Superior's Decision

In this model, the superior's instruction is payoff-irrelevant and unbinding. So, it only plays the role of conveying some information to the subordinate.

The superior's decision is not straightforward, because the superior has to take into account the subordinate's ability to infer from his instruction and the strategic aspect, in other words, the possibility that the subordinate behaves in the opposite way to his direction just for building up the reputation that he is able. Notice, however, that the superior with a very high v_c will be sure that v_s is also very high. Therefore, the superior will want the subordinate to choose $a = 1$ and the subordinate is very likely to choose $a = 1$ regardless of his ability. This common interest between the superior and the subordinate may enable the superior to transmit his private information v_c effectively through his recommendation.

For the following analysis, we will focus on the least-cost pooling equilibrium involving v_U in the continuation game.

Let $\tilde{v}_c$ be the equilibrium cutoff value of v_c suggesting $a = 0$ or 1. Also, let $\hat{v}_0$, $\hat{v}_1$ be the pooling cutoff values of the subordinate given that $v_c \leq \tilde{v}_c$ and $v_c > \tilde{v}_c$ respectively.

Suppose that the superior with $v_c \leq \tilde{v}_c$ recommends $a = 0$ and the one with $v_c > \tilde{v}_c$ recommends $a = 1$. The expected payoff of the superior with v_c when he recommends $a = 0$ is given by

$$\Delta_0(v_c) \equiv (1-\alpha)E(\theta_1; v_s > \hat{v}_0 \mid v_c) = (1-\alpha)E(\theta_1 \mid v_c, v_s > \hat{v}_0)\text{Prob}[v_s > \hat{v}_0 \mid v_c]. \quad (22)$$

If the superior recommends $a = 0$, he is expected to get a share of the realization of θ_1 only when the subordinate says no by choosing $a = 1$ which occurs if $v_s > \hat{v}_0$. On the other hand, his expected payoff when he suggests $a = 1$ is

$$\Delta_1(v_c) \equiv (1-\alpha)E(\theta_1; v_s > \hat{v}_1 \mid v_c) = (1-\alpha)E(\theta_1 \mid v_c, v_s > \hat{v}_1)\text{Prob}[v_s > \hat{v}_1 \mid v_c]. \quad (23)$$

Similarly, if the superior recommends $a = 1$, he is expected to get $(1 - \alpha)\theta_1$ only when the subordinate says yes by choosing $a = 1$.

Intuitively, if $\hat{v}_0$ is high enough, the superior will not have to be concerned about the possibility of disobey too much. So, a superior with small v_c who wants $a = 0$ to be implemented will find it optimal to recommend $a = 0$. Similarly, if $\hat{v}_1$ is very low, implying that the probability of obey is high, a superior with high v_c will prefer recommending $a = 1$ since $a = 1$ is what he actually wants to be chosen.

It is, in general, quite complicated to identify $\tilde{v}_c$ that makes the superior indifferent between recommending $a = 0$ and $a = 1$, i.e., $\Delta_0(\tilde{v}_c) = \Delta_1(\tilde{v}_c)$. However, the following theorem shows that $v_c = 0$ is the borderline value, if the pooling equilibrium in the subgame is symmetric in the sense $\hat{v}_1 < 0 < \hat{v}_0$ and $|\hat{v}_0| = |\hat{v}_1|$.

Theorem 5 There is a symmetric equilibrium in which the superior recommends $a = 0$ *if* $v_c \leq 0$ *and* $a = 1$ *if* $v_c > 0$.[11]

Proof. See the appendix.

This theorem says that the superior's suggestion can convey truthful information in equilibrium.

7 Discussion and Conclusion

This paper provides rationales for "yes men" and "no men," and attempts to explain why some behave as "yes men" and others as "no men" in an organization or a group. A person's tendency towards a "yes man" relies on his information and the accuracy of the information, and the tendency to be a "no man" depends on how much he cares about his reputation of being an expert. Also, in the light of a separating equilibrium obtained in Section 6, we can say that defiance can signal talent.

In this paper, we drew attention to the positive aspect of "yes men" and "no men," but the normative aspect also deserves to be addressed. In our model, it is obvious that behavior as a "no man" leads to inefficient decision-making in an organization, whereas behavior as a "yes man" who utilizes all the relevant information is encouraging, considering that information gathering is usually costly. In this view, it will be an important and promising issue to design an incentive scheme of an organization to minimize inefficiencies associated with "no men." As a related issue, we have not asked to whom

[11] In fact, for any equilibrium in a cheap talk game, there always exists a mirror equilibrium in which the messages in the original equilibrium are interpreted as the opposite. In other words, the superior recommends $a = 1$ if and only if $v_c < 0$ and the subordinate interprets $a = 1$ as $v_c < 0$ and $a = 0$ as $v_c \geq 0$. However, we will focus only on the intuitive equilibrium in which $a = 0$ ($a = 1$ resp.) is interpreted as $v_c \leq 0$ ($v_c > 0$ resp.).

responsibility should be allocated. Obviously, the allocation of responsibility is one important factor affecting the incentive of a person to obey the instructor. In fact, there are some studies in social psychology suggesting that individuals are less likely to obey and tend to follow their own judgment when they feel responsible for their own actions. (Hamilton [1978]; Tilker [1970]) However, if the superior makes the subordinate responsible only for the outcome resulting from his disobedience by a payment schedule incorporating a penalty and a bonus, the subordinate will be more likely to obey the superior as he is more risk averse. (See Kim and Ryu [1998].) Finally, a person may behave differently, depending on how much he cares about his reputation. Thus, the analysis for the model where the reputational concern, measured by k, is also the private information for the subordinate will be a challenging future research topic.

Appendix

A.1 Proof of Proposition 1

Lemma 2 $\theta_1 \mid v_c, v_s$ *is normally distributed with mean* $(h_c v_c + h_s v_s)/\Sigma$ *and precision* Σ*, where* $\Sigma \equiv h_1 + h_c + h_s$.

Proof. straightforward

Lemma 3 Let $\delta = h_s v_s/(h_1 + h_s)$, $\gamma = \sqrt{1/h_c + 1/(h_1 + h_s)}$, $\lambda = \phi(-\delta/\gamma)/\Phi(-\delta/\gamma)$ *where* $\phi(\cdot)$ *and* $\Phi(\cdot)$ *are the standard normal density function and distribution function respectively. Then, (i)* $E(\theta_1 \mid v_c \leq 0, v_s) = h_s v_s/(h_1 + h_s) - h_c\gamma\lambda/(h_1 + h_s + h_c)$ *(ii)* $Var(\theta_1 \mid v_c \leq 0, v_s) = 1/(h_1 + h_s + h_c) + h_c(1 + \delta\lambda/\gamma - \lambda^2)/((h_1 + h_s + h_c)(h_1 + h_s))$.

Proof. Since the conditional distribution of v_c given v_s is normal with mean δ and variance γ^2, we can represent the conditional distribution of v_c given v_s and $v_c < 0$ as that of $\delta + \gamma Z_1^*$, where Z_1^* is a standard normal random variable, truncated such that $Z_1^* < -\delta/\gamma$. Also, by Lemma 2, we can represent the conditional distribution of θ_1 given v_s and v_c as that of $h_s v_s/(h_1 + h_s + h_c) + h_c v_c/(h_1 + h_s + h_c) + \sqrt{Z_2/(h_1 + h_s + h_c)}$, where Z_2 is a standard normal random variable independent of Z_1^*. Finally, by combining the above two representations, we can represent the conditional distribution of θ_1 given v_s and $v_c \leq 0$ as that of $h_s v_s/(h_1 + h_s + h_c) + h_c(\delta + \gamma Z_1^*)/(h_1 +$

$h_s + h_c) + Z_2/\sqrt{h_1 + h_s + h_c}$. Now, the results in Lemma 3 follow from (i) $E(Z_1^*) = E(Z \mid Z < -\delta/\gamma) = -\lambda$ and (ii) $\mathrm{Var}(Z_1^*) = \mathrm{Var}(Z \mid Z < -\delta/\gamma) = 1 + \delta\lambda/\gamma - \lambda^2$.

Since v^0 must satisfy $E[\theta_1 \mid v_c \leq 0, v^0] = 0$, Lemma 3 implies that

$$h_s v^0/h_c = (h_1 + h_s)\gamma\lambda^0/(h_1 + h_s + h_c), \tag{A1}$$

where λ^0 is defined as the value of λ when $v_s = v^0$ is used in the definition of δ. Thus, the proof is immediate from equation (A1).

A.2 Proof of Lemma 1

(i) $\partial E[\theta_1 \mid v_c \leq 0, v_s, h_s]/\partial v_s = h_s/(h_1 + h_s) + [h_s/(h_1 + h_s)][h_c/(h_1 + h_s + h_c)]\lambda'(-\delta/\gamma) > 0$, if we notice that $\lambda'(x) < 0\ \forall x$ and that $|\lambda'(x)| < 1\ \forall x$. Proofs of the rest part of Lemma 1 are trivial.

A.3 Proof of Theorem 1

First, in a separating equilibrium, it is not possible that $\hat{v_H} > \hat{v_L}$. This is because a high-ability subordinate with any $v_s \in (\hat{v_L}, \hat{v_H}]$ will be better off by disobeying, as long as a low-ability subordinate with the same v_s prefers disobeying the instruction. Thus, it must be that $\hat{v_H} < \hat{v_L}$. In this equilibrium, the concept of the WPBE requires that the posterior belief be

$$\hat{\rho}(a(v_s)) = \begin{cases} 1 & \text{if } a(v_s) = 1 \text{ for } v_s \in (\hat{v_H}, \hat{v_L}] \\ 0 & \text{if } a(v_s) = 0 \text{ for } v_s \in (\hat{v_H}, \hat{v_L}] \\ \rho & \text{if } a(v_s) = 1 \text{ for } v_s \in (\hat{v_L}, \infty) \text{ or } a(v_s) = 0 \text{ for } v_s \in (-\infty, \hat{v_H}], \end{cases}$$

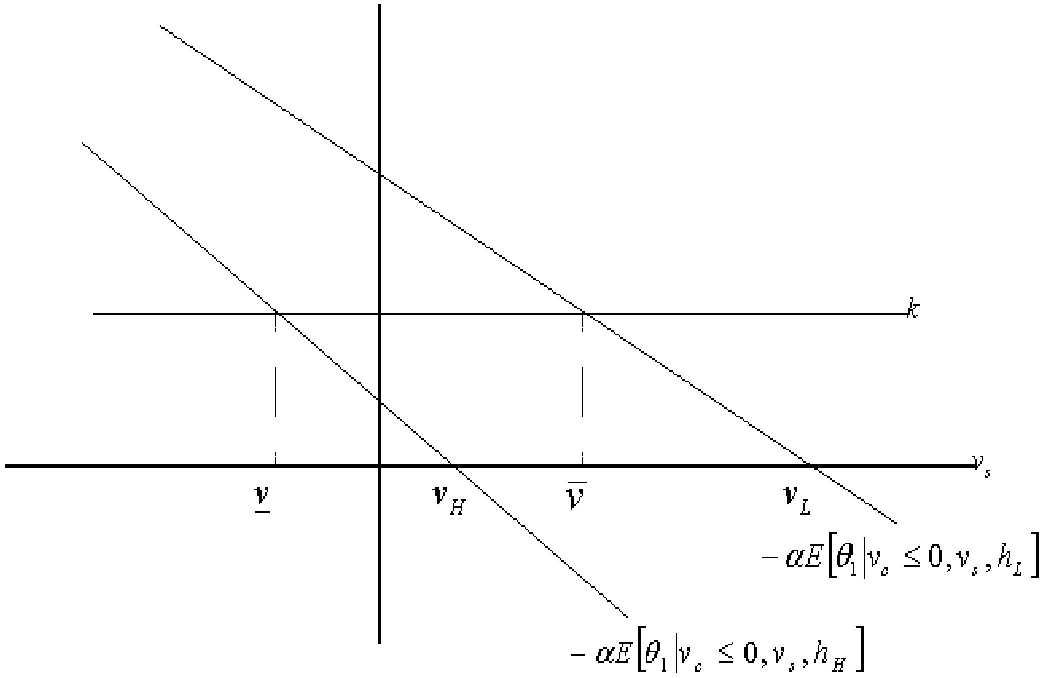
where $a(v_s)$ is the action taken by the subordinate with v_s and $\hat{\rho}(a(v_s))$ is the corresponding posterior belief. Also, by exploiting the arbitrariness of off-the-equilibrium-path beliefs, we assign the most pessimistic one, $\hat{\rho}(a(v_s)) = 0$, if $a(v_s) = 0$ for $v_s \in (\hat{v_L}, \infty)$, or $a(v_s) = 1$ for $v_s \in (-\infty, \hat{v_H}]$. The incentive compatibility conditions imply that, for any $v_s \in (\hat{v_H}, \hat{v_L}]$,

$$\alpha E[\theta_1 \mid v_c \leq 0, v_s, h_L] + k \leq 0, \tag{A2}$$

$$\alpha E[\theta_1 \mid v_c \leq 0, v_s, h_H] + k \geq 0. \tag{A3}$$

Let $\overline{v}$, $\underline{v}$ be v_s satisfying (A2), (A3) respectively with equality. Then, in order for $(\hat{v_L}, \hat{v_H})$ to be equilibrium cutoff values, it must be that $[\hat{v_H}, \hat{v_L}] \subset [\underline{v}, \overline{v}]$.

Figure A1
Nonexistence of Separating Equilibrium



On the other hand, for all $v_s \in (\hat{v_L}, \infty)$, both types of subordinate, in particular, the low-ability one must say no. In other words, it must be that

$$(A4) \qquad \alpha E[\theta_1 \mid v_c \leq 0, v_s, h_L] + k\rho \geq 0.$$

Since $E[\theta_1 \mid v_c \leq 0, v_s, h_L]$ is continuous in v_s, there always exists some $\epsilon > 0$ such that no $v_s \in (\hat{v_L}, \hat{v_L} + \epsilon)$ can satisfy inequality (A4), which implies that $\hat{v_L}$ cannot be the equilibrium cutoff value of the low-ability subordinate.

A.4 Proof of Theorem 2

Take any $\hat{v} \in [\underline{v^p}, \overline{v^p}]$. It is straightforward to see that any $v_s > \hat{v}$ satisfies (19) and that any $v_s \leq \hat{v}$ satisfies (20). Also, the possibility that $\underline{v^p} < 0$ is implied by equation (19) when k or ρ is very large.

A.5 Proof of Theorem 4

Take $\hat{v}_H = \underline{v}$ and $\hat{v}_L = \overline{v}$. Then, it is sufficient to find the equilibrium values of $r(v_s)$ and pin down the posterior beliefs to support this as an equilibrium.

First, the low-ability subordinate must be indifferent between $a = 0$ and $a = 1$ for any $v_s \in (\hat{v_L}, v_\rho]$, given $\hat{\rho}(1; v_s)$ where $\hat{\rho}(1; v_s)$ is the posterior probability that the subordinate is of high ability given that the colleagues observe the subordinate with $v_s \in (\hat{v_L}, v_\rho]$ disobeying. Thus, we have

$$\alpha E[\theta_1 \mid v_c \leq 0, v_s, h_L] + k\hat{\rho}(1; v_s) = 0. \tag{A5}$$

On the other hand, $\hat{\rho}(1; v_s)$ must be updated according to the following rule

$$\hat{\rho}(1; v_s) = \frac{\rho}{\rho + (1-\rho)r(v_s)}. \tag{A6}$$

Combining these, we get

$$r(v_s) = -[1 + \frac{k}{\alpha E(\theta_1 \mid v_c \leq 0, v_s, h_L)}]\frac{\rho}{1-\rho}, \tag{A7}$$

$$\hat{\rho}(1; v_s) = -\frac{\alpha E[\theta_1 \mid v_c \leq 0, v_s, h_L]}{k}. \tag{A8}$$

Notice that $\hat{\rho}(1; \overline{v}) = 1$, $\hat{\rho}(1; v_\rho) = \rho$, $r(\overline{v}) = 0$ and $r(v_\rho) = 1$.

A.6 Proof of Theorem 5

Let $\Delta(w) = E(\theta_1; v_s > w \mid v_c)$. Then,

$$\begin{aligned}
\Delta(w) &= \int_{-\infty}^{\infty}\int_{w}^{\infty} \theta_1 f(\theta_1, v_s \mid v_c) dv_s d\theta_1 \\
&= \int_{-\infty}^{\infty}\int_{w}^{\infty} \theta_1 \frac{f(\theta_1, v_s, v_c)}{f_c(v_c)} dv_s d\theta_1 \\
&= \int_{-\infty}^{\infty}\int_{w}^{\infty} \theta_1 \frac{\rho f(\theta_1, v_s, v_c \mid h_H) + (1-\rho) f(\theta_1, v_s, v_c \mid h_L)}{f_c(v_c)} dv_s d\theta_1,
\end{aligned}$$

where $f(\theta_1, v_s, v_c)$ is the joint density function of θ_1, v_s, v_c and f_c is the marginal density function of v_c.

Now, given either $h_s = h_L$ or $h_s = h_H$, $f(\theta_1, v_s, v_c)/f_c(v_c)$, as a conditional joint density function of θ_1 and v_s given v_c, follows a normal distribution. To characterize this conditional density function, we borrow the following lemma from the statistics literature.

Lemma 4 If $\begin{pmatrix} y \\ x \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}, \begin{pmatrix} \Sigma_{yy}\Sigma_{yx} \\ \Sigma_{xy}\Sigma_{xx} \end{pmatrix} \right)$, *we have* $y \mid x \sim N(\mu_{y|x}, \Sigma_{yy|xx})$, *where* $\mu_{y|x} = \mu_y + \Sigma_{yx}\Sigma_{xx}^{-1}(x - \mu_x)$, $\Sigma_{yy|xx} = \Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}$.

To use the above result, let $x = v_c$, $y = (\theta_1, v_s)^T$. Notice that

$$\begin{pmatrix} y \\ x \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 + \sigma_s^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2 + \sigma_c^2 \end{pmatrix} \right),$$

where $\sigma_1^2 = 1/h_1$, $\sigma_i^2 = 1/h_i$, $i = c, s$. Then, it follows that the conditional mean and variance are

$$\mu_{y|x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \sigma_1^2 \\ \sigma_1^2 \end{pmatrix} \frac{v_c}{\sigma_1^2 + \sigma_c^2} = \begin{pmatrix} pv_c \\ pv_c \end{pmatrix},$$

$$\Sigma_{yy|xx} = \begin{pmatrix} \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 + h_s^2 \end{pmatrix} - \begin{pmatrix} \sigma_1^2 \\ \sigma_1^2 \end{pmatrix} \frac{1}{\sigma_1^2 + \sigma_c^2} \begin{pmatrix} \sigma_1^2 & \sigma_1^2 \end{pmatrix} = \sigma_1^2 \begin{pmatrix} 1-p & 1-p \\ 1-p & 1-p+\frac{\sigma_s^2}{\sigma_1^2} \end{pmatrix},$$

where $p = \sigma_1^2/(\sigma_1^2 + \sigma_c^2)$. Therefore, $\Delta(w)$ can be written as

$$\Delta(w) = \rho E(W_1; W_2 > w \mid h_H) + (1-\rho)E(W_1; W_2 > w \mid h_L),$$

where $(W_1, W_2)^T \mid h_s \sim N(\mu_{y|x}, \Sigma_{yy|xx} \mid h_s)$.

Now, let us use the Choleski decomposition.

Lemma 5 (Choleski Decomposition) Let Z_1 and Z_2 be two i.i.d. standard normal random variables. Then, there exist constants c_{11}, c_{12} and c_{21} such that $\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \stackrel{\mathrm{D}}{=} \mu_{y|x} + \begin{pmatrix} c_{11}Z_1 + c_{12}Z_2 \\ c_{21}Z_1 \end{pmatrix}$, *where '$\stackrel{\mathrm{D}}{=}$' denotes that both hand sides follow the same distribution.*

Then, we have $\Delta(w) = \rho E(pv_c + c_{11}Z_1 + c_{12}Z_2; pv_c + c_{21}Z_1 > w \mid h_H) + (1-\rho)E(pv_c + c_{11}Z_1 + c_{12}Z_2; pv_c + c_{21}Z_1 > w \mid h_L) = \rho pv_c\Phi(\frac{pv_c - w}{c_{21}} \mid h_H) + \rho c_{11}\phi(\frac{pv_c - w}{c_{21}} \mid h_H) + (1-\rho)pv_c\Phi(\frac{pv_c - w}{c_{21}} \mid h_L) + (1-\rho)c_{11}\phi(\frac{pv_c - w}{c_{21}} \mid h_L)$.

Now, it suffices to show that, if $\tilde{v}_c = 0$, $\Delta(\hat{v}_0) = \Delta(\hat{v}_1)$ when $\hat{v}_1 < 0 < \hat{v}_0$ and $|\hat{v}_0| = |\hat{v}_1|$. When $v_c = 0$, $\Delta(w) = \rho c_{11}\phi(-\frac{w}{c_{21}} \mid h_H) + (1-\rho)c_{11}\phi(-\frac{w}{c_{21}} \mid h_L)$, which is symmetric about $w = 0$. This completes the proof.

References

Asch, S. [1956], "Studies of Independence and Conformity: A Minority of One against a Unanimous Majority," *Psychological Monographs*, 70, 9.

Banerjee, A. [1992], "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797-817.

Banks, J., and J. Sobel [1987], "Equilibrium Selection in Signaling Games," *Econometrica*, 55, 647-661.

Bikhchandani, S., D. Hirshleifer, and I. Welch [1992], "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992-1026.

Crawford, V., and J. Sobel [1982], "Strategic Information Transmission," *Econometrica*, 50, 1431-1451.

Croner, M., and R. Willis [1961], "Perceived Differences in Task Competence and Asymmetry of Dyadic Influence," *Journal of Abnormal and Social Psychology*, 62, 705-708.

Crutchfield, R. A. [1955], "Conformity and Character," *American Psychologist*, 10, 191-198.

Hamilton, V. L. [1978], "Obedience and Responsibility: A Jury Simulation," *Journal of Personality and Social Psychology*, 36, 126-146.

Kim, J., and K. Ryu [1998], "Responsibility and Risk Aversion in Obedience to Authority," Mimeo, Department of Economics, Dongguk University, Seoul, Korea.

Levy, G. [2000], "Strategic Consultation of Advisers in the Presence of Career Concerns," London School of Economics STICERD Discussion Paper TE/00/404, London, UK.

Marlowe, D., and K. J. Gergen [1970], "Personality and Social Behavior," pp. 1-76, in: K. J. Gergen and D. Marlowe (eds.), *Personality and Social Behavior*, MA: Addison-Wesley.

Mischel, W. [1968], *Personality and Assessment*, New York: Wiley.

Ottaviani. M., and P. Sørensen [2001], "Information Aggregation in Debate: Who Should Speak First?," *Journal of Public Economics*, 81, 393-421

Prendergast, C. [1993], "A Theory of 'Yes Men'," *American Economic Review*, 83, 757-770.

Prendergast, C., and L. Stole [1996], "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," *Journal of Political Economy*, 104, 1105-1134.

Scharfstein, D., and J. Stein [1990], "Herd Behavior and Investment," *American Economic Review*, 80, 465-479.

Tilker, H. A. [1970], "Socially Responsible Behavior as a Function of Observer Responsibility and Victim Feedback," *Journal of Personality and Social Psychology*, 14, 95-100.

Trueman, B. [1994], "Analyst Forecasts and Herding Behavior," *Review of Financial Studies*, 7, 97-124.

Jeong-Yoo Kim
Department of Economics
Dongguk University
3-26, Pildong, Chungku
Seoul 100-715, Korea
Email: jyookim@dongguk.edu

Keunkwan Ryu
Department of Economics
Seoul National University
Seoul 151-742, Korea
Email: ryu@snu.ac.kr