Intrinsic Motivation in Collectivist Societies

Frank Fabian
Peter-J. Jost

Discussion Paper No. 09-03
Intrinsic Motivation in Collectivist Societies

Fabian Frank and Peter-J. Jost

Abstract

In recent years, intrinsic motivation has been an important subject of personnel economics. The cross-cultural psychology literature suggests that motivational effects vary across cultures. Cultural variations in intrinsic motivation have not been analyzed economically, yet. We combine the frameworks of Bénabou and Tirole (2003) and Sliwka (2007) to show that in collectivist societies, intrinsic motivation is higher and the use of contingent rewards lower than in individualistic ones. Furthermore, we find that intrinsic motivation is lower in societies with a homogeneous work ethos.

1 Introduction

In several interviews in Arab and Chinese firms, we found that companies in collectivist societies emphasize immaterial incentives and the strengthening of employees' intrinsic motivation to increase employee effort. Both the Arab and the Chinese culture have been identified as collectivist in Hofstede (2001) and Gelfand et. al. (2004). For example, National Bank of Abu Dhabi, the second biggest bank by assets of the United Arab Emirates (http://archive.gulfnews.com/indepth/dubaibankmerger/more_stories/10109387.html, accessed on June 30, 2008), focuses on immaterial incentives. In addition to a monetary incentive scheme, the bank provides a wide range of immaterial incentives such as reward ceremonies for long-standing employees, a control system of line managers to ensure appreciation of well performing employees and the possibility that the bank employs the wife or the son of an employee who has been in the bank for at least 15 years. The National Petrochemical Industrial Company of Saudi Arabia considers the strengthening of its employees' intrinsic motivation to be decisive. Showing respect by inviting poor employees and their families to health and sport clubs, by introducing joint company activities and by sending letters of thanks from the president to individual employees caused a significant rise in motivation. The Dubai auditing firm Farahat & Co. experienced that European employees focus on monetary rewards while Arab and Asian employees depend

*WHU - Otto Beisheim School of Management, Burgplatz 2, 56179 Vallendar, Germany, Email: pjjost@whu.edu
on the respect which the employer shows to them. Specifically, Europeans on the one hand were not prepared to wait when requesting a salary increase. Arabs and Asians on the other hand were content with delaying the salary increase. However, Arab and Asian employees need to be shown respect in a much more pronounced way. In order to show his respect, the company owner offers food after the end of work, cares about the well-being of his employees’ families and establishes long-term personal relationships. Changshu Automotive Trim Corporation (CAIP) of China is faced with the problem that employees in some regions work less hard than employees in other regions of China. To solve the issue, bonuses have been tried but a full improvement turned out to be only possible with personal talks, particularly with the employees’ parents and families.

Starting from these indications from company practice, we will provide an economic analysis of intrinsic motivation in collectivist societies. Intrinsic motivation is given when "the rewards are inherent in the activity, and even though there may be secondary gains, the primary motivators are spontaneous, internal experiences that accompany the behavior." (Deci and Ryan, 1985, p. 11). By contrast, "extrinsic motivation refers to behavior where the reason for doing it is something other than an interest in the activity itself." (Deci and Ryan, 1985, p. 35).

The economic literature has focused particularly on the negative effect of monetary incentives on intrinsic motivation. Frey (1997) as well as Frey and Oberholzer-Gee (1997) show that extrinsic incentives can crowd out intrinsic motivation. Frey and Jegen (2001) cite a number of empirical studies which find a crowding-out effect of external interventions via monetary incentives or punishments. The model of Bénabou and Tirole (2003) points out that setting a contingent reward can be a negative signal about an agent’s cost of effort. Bénabou and Tirole (2006) emphasize that rewards create doubt about the true motive of prosocial behavior and therefore result in a crowding-out effect. Sliwka (2007) explains the crowding-out effect in a situation in which the principal trusts or controls depending on her beliefs about the distribution of agent types in the population.

A general conclusion from these studies is that when employees have high levels of intrinsic motivation, setting extrinsic incentives might have particularly negative effects on exerted effort levels. Following the study by Erez (1994) on cross-cultural psychology, however, the nature of intrinsic motivation varies across cultures.\footnote{See also Kurman (2001) who studies self-regulation strategies across cultures and states that intrinsic motivation might be one aspect leading to different levels of self-regulation. However, not all researchers in the cross-cultural literature follow the hypothesis that there exist cultural differences in intrinsic motivation. Sheldon et.al. (2004) for example believe that people can be intrinsically motivated in almost any context and do not include this hypothesis in their cross-cultural study.} In particular, as Gelfand et.al. (2004) pointed out, the nature of motivation varies in individualistic and collectivist societies. The terms individualism and collectivism thereby refer to the extent to which people are autonomous individuals or embedded in their groups. Hofstede (2001) defines
"Individualism stands for a society in which the ties between individuals are loose: Everyone is expected to look after him/herself and her/his immediate family only. Collectivism stands for a society in which people from birth onwards are integrated into strong, cohesive in-groups, which throughout people’s lifetime continue to protect them in exchange for unquestioning loyalty." (p. 225).

In collectivist societies, fulfilling duties and obligations as well as contributing to the group are central components of motivation. Hofstede (2001) emphasizes that organization members should therefore have a higher emotional dependence on their organizations. He points out that we can expect more "moral" involvement with the organization in collectivist societies and more "calculative" involvement in individualistic societies (p. 212). Accordingly, employers might rely more on intrinsic motivation in collectivist and on extrinsic incentives in individualistic societies. This observation is confirmed by several studies on cross-cultural psychology. Gelfand et.al. (2004) point out that employees in individualistic cultures are more motivated by reward contingent on performance. Consequently, companies in individualistic societies distribute rewards differentially, and contingent on performance, see Erez (1994). Hempel (1998) writes that employees in collectivist societies should resist flexible benefit programs. Moreover, in collectivist societies, organizations are less likely to focus on rational exchanges with their members, see Gelfand et.al. (2004). Their empirical studies show that contingent rewards are not related to employee performance and that they are used to a smaller extent in collectivist societies such as Mexico or India.

Building on these insights, we propose an economic analysis of cross-cultural differences on intrinsic motivation. The basic framework for this analysis is similar to Bénabou and Tirole (2003) and Sliwka (2007). In particular, we consider a principal-agent relationship in which the principal offers a performance-based bonus to induce the agent to perform a task. The agent can be one of three types: a workaholic, lazy or conformist agent. Workaholic and lazy agents differ in their work ethos, expressed in low or high cost of effort. Conformists, however, are influenced by social norms so that their work ethos depends on the median work ethos in the society. The distribution of types in the population is uncertain. While each agent knows his own type, only the principal observes the

---

2Some parts of the cross-cultural psychology literature consider the concepts of individualism and collectivism as an oversimplified view of cultural differences, see Baskerville (2003), Baskerville-Morley (2005) or Efferin and Hopper (2007). Particularly, Hofstede is criticized for proposing socio-economic rather than cultural dimensions, see Baskerville (2003). Authors from this branch of the literature suggest that cross-cultural analyses should be based on individual values of particular cultures, e.g. Efferin and Hopper (2007) or use constructs from current social anthropology, e.g. Baskerville (2003). As Hofstede (2003) points out, however, whether collectivism is a cultural or a socio-economic dimension, it still describes an important characteristic which organizations need to consider when they move across borders. Consequently, we use individualism versus collectivism as a construct which allows us to shed light on fundamental differences in employee motivation across societies.
distribution of agent types because of her experience with current and former employees.

In this framework, we are interested in a separating equilibrium in which the principal uses the size of the bonus to signal her private information about the distribution of work ethos in the population. We show that the principal chooses a low bonus signaling a social norm of low cost of effort when the number of workaholics and the fraction of conformists are high. In this case she can credibly signal her private information such that conformists adopt the work ethos of the workaholics and suffer low effort cost. Using the definition of intrinsic motivation proposed by Bénabou and Tirole (2003), a higher cost of effort reduces the intrinsic motivation of an agent. Accordingly, when the conformists' cost of effort decreases following the credible signal of a low bonus, intrinsic motivation increases. Following the paradigm of individualism and collectivism we identify collectivism with an environment in which more employees conform to the social norm than in case of individualism. Hence, when the fraction of conformist agents in the population is high (as in a collectivist society), intrinsic motivation is higher and the use of contingent rewards lower than in individualistic ones. We also find that intrinsic motivation is reduced when the differences between the work ethos of different agents are small. In societies whose work ethos is homogeneous intrinsic motivation is therefore lower than in societies with high differences in work ethos.

The paper proceeds as follows. In section 2, we set up our model. Section 3 contains the main analysis of our model. We will sum up our findings in section 4.

## 2 The model

We use the framework of Bénabou and Tirole (2003, p. 496ff) to model the interactions between a principal and an agent: A risk-neutral principal ("she") employs an agent ("he"). The agent chooses whether to exert effort or not, i.e., \( e \in \{0, 1\} \). If he does not work, \( e = 0 \), his cost of effort is normalized to zero. When choosing high effort, \( e = 1 \), his effort cost \( c \) is either low or high, \( c \in \{c_L, c_H\} \) with \( 0 < c_L < c_H \). The actual effort cost reflects the work ethos of the agent. We call an agent workaholic if he has low cost of effort \( c_L \), whereas the agent is called a lazy agent if he has high cost of effort \( c_H \). Since the work ethos of these two types is fixed from the beginning, we call these agents steadfast.

We follow the idea of Sliwka (2007) and assume that besides being a workaholic or a lazy agent, there is a third group of agents called conformists. A conformist is someone who is uncertain about the "appropriate" behavior in a certain situation and therefore is influenced by social norms. If, for instance, a conformist works a lot, he will suffer from too much work only if he believes that many others would also feel bad about the workload. We model this by

---

3This is consistent with the finding in the literature that collectivism is linked with conformity, see Gelfand et.al. (2004).
assuming that there is uncertainty about the fraction $\phi$ of workaholic agents among the steadfast agents. For simplicity we assume that $\phi \in \{\phi_L, \phi_H\}$ with $\phi_L < \frac{1}{2} < \phi_H$ with a common prior expectation of $\phi$ lower than $\frac{1}{2}$. Hence, the cost of effort of a conformist is equal to $c_L$ if he believes that the median work ethos is to be workaholic (i.e. if his conditional expectation on $\phi$ is larger than $\frac{1}{2}$) and equal to $c_H$ otherwise. The fraction of conformists in the population has expected value $\eta$ according to the common prior expectation. According to Sliwka (2007) and Bénabou and Tirole (2003), as an employer, the principal will typically have learned more from the behavior of previous or other current employees. For simplicity, we assume that she learns the fraction of workaholic agents $\phi$.

The verifiable outcome of the agent’s work can be either good or bad. It is good with probability $\theta$, where $\theta \in (0, 1]$ is identical for all types of agent and is common knowledge. If the outcome is good it yields a direct payoff $V$ to the agent and a profit $W$ to the principal, with $V < c_L$ and $W > \frac{\theta}{2} - V$. In case the outcome is bad, both gross payoffs equal 0.

In order to induce the agent to exert effort, the principal offers a performance-based reward or "bonus" $b$, $b \in [0, W]$. The principal can only set one performance-based bonus size for all agents. The agent’s net benefit in case of a good outcome is $V + b$, the principal’s is $W - b$. The resulting utility function of an agent equals $U = \theta(V + b) - c$ if he chooses $e = 1$, and the principal’s payoff from the agent in this case is $\Pi = \theta(W - b)$.

Note, that a steadfast agent who knows his cost $c \in \{c_L, c_H\}$ of effort has an intrinsic motivation of $(\theta V - c)$ and an extrinsic motivation of $\theta b$ to exert high effort. When, however, an agent is unsure about his cost of effort, he updates his beliefs using the bonus offered. Let $\hat{c}(b) = E[c | b]$ denote the agent’s expectation about the cost of effort given the principal offered bonus $b$. Then his intrinsic motivation is $(\theta V - \hat{c}(b))$.

The timing of the game is as follows: In Stage 1 the principal learns her private signal about $\phi$ and chooses the size of the bonus $b$. After observing the bonus chosen by the principal, the agent chooses in Stage 2 an effort level $e$. The principal’s payoff is realized in Stage 3 and the payment to the agent is made.

3 Collectivism and intrinsic motivation

The game between the principal and the agent is a signaling game, as the principal’s choice may reveal her private information about the fraction of workaholic agents $\phi$. In the following, we are interested in separating equilibria in which the principal will set a high bonus $b_H$ if her private information is bad, that is $\phi = \phi_L$, and a low bonus $b_L$ in the opposite case if her private information is good, that is $\phi = \phi_H$. The choice of a low or high bonus then signals her private

---

4Since the probability of success $\theta$ is common knowledge and the principal and the agent are both risk-neutral, a situation in which the principal observes the agent’s effort is equivalent to a situation in which she can observe only outcome: The reward contingent on effort is identical to the one contingent on outcome.
information and is informative in the sense that it affects the preferences of a conformist given that he forms a conditional expectation $E[\phi|b]$ based on the principal’s signal.

Before we check that the principal actually has an incentive to follow this strategy when a conformist agent believes that the bonus is a credible signal, consider the behavior of a steadfast agent who knows his cost of effort are $c$. He chooses effort $e = 1$ if and only if

$$\theta(V + b) \geq c. \quad (1)$$

The interesting case then is when for the bonuses $b_H$ and $b_L$ the conditions $\theta(V + b_L) < c_H$, $\theta(V + b_H) \geq c_H$, and $\theta(V + b_L) \geq c_L$ are fulfilled: To a steadfast lazy agent who has perfect information that his cost of effort equals $c$, he chooses effort $e = 1$ if and only if

$$b_H = \frac{c_H}{\theta} - V \quad \text{and} \quad b_L = \frac{c_L}{\theta} - V.$$ 

If $\phi$ denotes the principal’s subjective probability that a steadfast agent is a workaholic, her overall payoff when offering a high bonus $b_H$ reads as

$$\Pi(b_H) = \eta \theta(W - b_H) + (1 - \eta)(\phi \theta(W - b_H) + (1 - \phi)\theta(W - b_H)) \quad (2)$$

$$= \theta(W - b_H).$$

If she offers a low bonus $b_L$, her expected payoff equals

$$\Pi(b_L) = \eta \theta(W - b_L) + (1 - \eta)\phi \theta(W - b_L) \quad (3)$$

$$= \theta(W - b_L).$$

Comparing these two equations we find that $\Pi(b_H)$ is strictly higher than $\Pi(b_L)$ whenever

$$\eta < \hat{\eta}(\phi) = 1 - \frac{b_H - b_L}{(W - b_L)(1 - \phi)} \quad (4)$$

Proposition 1 There exists a separating equilibrium in which the principal pays a low bonus $b_L$ after she has received a good signal and a high bonus $b_H$ after the bad signal if and only if the fraction of conformists is $\eta \in \left[\max\left\{0, \hat{\eta}(\phi_H)\right\}, \hat{\eta}(\phi_L)\right]$ and $\phi_L < \min\left\{\frac{1}{2}, \frac{\theta(W + V) - c_H}{\theta(W + V) - c_L}\right\}$.

Proof. Since $W > b_H > b_L$ the critical value $\hat{\eta}(\phi)$ is decreasing in $\phi$ and strictly smaller than 1. That is, $\hat{\eta}(\phi_H) < \hat{\eta}(\phi_L) < 1$. Note that, since $b_H - b_L = \frac{c_L}{\theta} - V$, $b_H > b_L > 0$ since $c_H > c_L$ and $c_L > V$.

Note, that the principal’s profit is always positive independent of the agent’s type, since $\theta(W + V) > c_H$ by assumption.

5 Note, that the agent’s participation constraint is always satisfied since, by assumption, the bonus cannot be negative.

6 Note, that $b_H > b_L > 0$ since $c_H > c_L$ and $c_L > V$.

7 Note, that the principal’s profit is always positive independent of the agent’s type, since $\theta(W + V) > c_H$ by assumption.
\( b_H - W < \phi(b_L - W) \)
\( \phi < \frac{W - b_H}{W - b_L} \).

Hence, \( \hat{\eta}(\phi_L) > 0 \) if \( \phi_L < \frac{\theta(W + V) - c_L}{\theta(W + V) - c_L} \). To see that the above conditions actually ensure a separating equilibrium note that as \( \phi_H > \frac{1}{2} > \phi_L \) the conformists will indeed be workaholics if and only if the principal offers \( b_L \). Hence, such an equilibrium will exist if she only prefers to offer \( b_L \) after the good signal. This will be the case whenever \( \eta \in [\max \{0, \hat{\eta}(\phi_H)\}, \hat{\eta}(\phi_L)] \). The condition
\( \phi_L < \min \left\{ \frac{1}{2}, \frac{\theta(W + V) - c_L}{\theta(W + V) - c_L} \right\} \) then ensures that this set is non-empty.8

A separating equilibrium even exists when the fraction of conformists \( \eta \) is small. Indeed, for \( \eta = 0 \), paying a low bonus \( b_L \) leads to higher (lower) profits for a principal than a high bonus \( b_H \) when she observed a good (bad) signal if and only if
\[
\phi_H \theta(W - b_H) + (1 - \phi_H)\theta(W - b_H) < \phi_H \theta(W - b_L) \tag{5}
\]
\[
\phi_L \theta(W - b_H) + (1 - \phi_L)\theta(W - b_H) > \phi_L \theta(W - b_L). \tag{6}
\]

Both inequalities are satisfied as long as9
\[
0 < \phi_L < \frac{1}{2} < \phi_H < 1 \quad \text{and} \quad \phi_L < \frac{W - b_H}{W - b_L}. \tag{7}
\]

On the other hand, if the fraction of conformists \( \eta \) is high, no separating equilibrium exists. If, for example, \( \eta = 1 \), calculations show that the principal always prefers to pay a low bonus \( b_L \) since
\[
\Pi(b_H) = \theta(W - b_H) < \Pi(b_L) = \theta(W - b_L). \tag{8}
\]

The intuition is straightforward: If the fraction of conformists is sufficiently high, the principal has an incentive to signal a high share of workaholics in order to pay a low bonus to conformists irrespective of the number of workaholics. Consequently, the principal’s signal is not credible any more for very high values of \( \eta \).10

8 Note that \( \frac{\theta(W + V) - c_L}{\theta(W + V) - c_L} \) is the higher the lower the difference between \( c_H \) and \( c_L \).

9 Note that \( 0 < \frac{W - b_H}{W - b_L} < 1 \).

10 For \( \eta > \hat{\eta}(\phi_L) \), there is a pooling equilibrium with \( b_H \). If the principal paid \( b_L \) for any \( \eta > \hat{\eta}(\phi_L) \), conformists would still turn into lazy agents since they would not believe her signal. She would then lose the efforts of the majority of agents: the steadfast lazy and the conformists turned lazy. These agent types are in the majority since \( \phi = \phi_L < \frac{1}{2} \). Consequently, the principal cannot benefit from paying \( b_L \) for very high numbers of conformists. See also the Appendix for the existence of further equilibria.
If we disregard extreme values \( \eta \), the proposition shows the principal can credibly use a performance-based reward to signal her private information about the distribution of work ethos in the population. The conformists will indeed adopt the work ethos of the workaholics since only the principal after receiving a good signal offers a low bonus. In this case, conformists suffer low effort cost. The following two corollaries are immediate consequences of this result:

**Corollary 2** In collectivist societies with a high share of conformists, intrinsic motivation is higher and the use of contingent rewards lower than in individualistic ones.

Using the proposition above, the conformists’ cost of effort decreases following the credible signal of a low bonus. That is, only when the principal pays a low bonus \( b_L \), conformists will have cost of effort \( c_L \). She is more likely to pay \( b_L \) when there are more conformists\(^{11} \). Since a higher cost of effort reduces the intrinsic motivation of an agent, intrinsic motivation increases. Identifying collectivism with a population in which more employees conform to the social norm than in case of individualism, the fraction of conformist agents in collectivist societies is higher than in individualistic ones. As a consequence, intrinsic motivation is higher and the use of contingent rewards lower in collectivist societies than in individualistic ones.

**Corollary 3** In homogeneous societies with a low difference in work ethos between workaholics and lazy agents, intrinsic motivation is lower and the use of contingent rewards more widespread than in societies with high differences in the work ethos of individual members.

In the proof above, we have shown that the cut-off value \( \tilde{\eta}(\phi) \) in the separating equilibrium is the higher the smaller the difference \( (c_H - c_L) \) in work ethos. Furthermore, the maximum \( \phi \) for which the principal’s signal is credible is the lower the smaller the difference \( (c_H - c_L) \) in work ethos. The principal will choose \( b_H \) more often in homogeneous societies and the conformists will adopt the low work ethos and high cost of effort \( c_H \) more frequently. For this reason, intrinsic motivation as defined in section 2 is smaller in homogeneous societies and the use of contingent rewards is higher than in societies with high differences in the work ethos of individual members.

### 4 Conclusion

Our model shows that intrinsic motivation is higher in collectivist societies. At the same time, the use of contingent rewards should be lower. Furthermore, we find that intrinsic motivation is lower in homogeneous societies. According to the model, the mechanism for this phenomenon is based on a signal about the employees’ work ethos. When there are steadfast workaholic, steadfast lazy and

\(^{11}\text{This is true if we disregard very high values of } \eta. \text{ Since in reality, no society will consist exclusively of conformists, we can in fact ignore extreme values of } \eta.\)
conformist agents and the principal has private information about their distribution, she may use the size of her bonus payment to signal the work ethos of the majority of steadfast agents. In case the signal is credible, the conformists will take on a work ethos corresponding to the principal’s bonus payment. Since the principal is more likely to set the low bonus for a higher number of conformists, in collectivist societies where many conformists are present, the principal will therefore pay the low bonus more often. Consequently, the conformists adopt the low cost of effort and their intrinsic motivation increases. The workings of the model are limited to medium numbers of conformists. When there are too many conformists, the principal’s signal is no longer credible. However, we can expect that even collectivist societies generally do not consist almost exclusively of conformists.

Our model has important implications for multinational companies. The use of a uniform incentive scheme for businesses in fundamentally different cultures may negatively affect employee motivation. For example, high reliance on contingent rewards to motivate employees in collectivist societies may reduce the employees’ intrinsic motivation and therefore decrease overall effort levels.

The findings of the model are consistent with some parts of the cross-cultural psychology literature. Although without formal analysis, this literature provides numerous indications that motivation is a phenomenon which varies across cultures, see Gelfand et. al. (2004). In particular, we could confirm the finding of the psychology literature, e.g. Erez (1994) or Gelfand et.al. (2004), that contingent reward should be used more frequently in individualistic societies than in collectivist ones. In order to further evaluate the result of the model, empirical studies of contingent rewards and intrinsic motivation across cultures would be beneficial. In the interviews in Chinese and Arab companies, we could already find certain indications that the use of contingent reward may indeed be lower in collectivist cultures.

5 Appendix

5.1 Further equilibria

5.1.1 Pooling equilibria for \( \phi_L > \frac{W - b_L}{\theta L} \):

In case \( \phi_L > \frac{W - b_L}{\theta L} \), \( \hat{\eta}(\phi_L) \) is smaller than zero and the principal would pay \( b_L \) irrespective of her private information. Consequently, her signal is no longer credible. For \( E[\phi] < \frac{1}{2} \) the conformists would turn into lazy agents since the principal’s policy does not have any signaling value. Taking this fact into consideration, the principal is confronted with the payoff functions

\[
\Pi(b_H) = \eta \theta (W - b_H) + (1 - \eta)(\phi \theta (W - b_H)) + (1 - \phi) \theta (W - b_H) \quad (9)
\]

and

\[
\Pi(b_L) = (1 - \eta) \phi \theta (W - b_L). \quad (10)
\]
The principal pays \( b_L \) if
\[
(W - b_H) < (1 - \eta)\phi(W - b_L)
\] (11)
or
\[
\eta < \hat{\eta}_P(\phi) = 1 - \frac{W - b_H}{\phi(W - b_L)}
\] (12)
She chooses \( b_L \) the more often the higher the number of workaholics \( \phi \). When \( \phi > \frac{W - b_H}{W - b_L} \), \( \hat{\eta}_P(\phi) > 0 \) while \( \hat{\eta}(\phi) < 0 \). However, as soon as she could get conformists to turn into supporters, her payoff functions would change and her signal following \( \hat{\eta}_P(\phi) \) cannot be credible. For \( \phi_L = \frac{W - b_H}{W - b_L} \),
\[
\hat{\eta}_P(\phi_L) = 1 - \frac{W - b_H}{\frac{W - b_H}{W - b_L}(W - b_L)} = 0.
\] (13)
In case \( \phi_L = \frac{W - b_H}{W - b_L} \), the minimum value of the right hand side is reached:
\[
(W - b_H) = (1 - \eta)\frac{W - b_H}{W - b_L}(W - b_L)
\] (14)
and \( b_H \) is optimal. If \( \phi_L < \frac{W - b_H}{W - b_L} \), the original signal becomes credible again and \( \hat{\eta}_P(\phi) < 0 \) while \( \hat{\eta}(\phi) > 0 \).

For very low numbers of conformists \( \eta < \hat{\eta}_P(\phi_L) \), there is a pooling equilibrium with \( b_L \). If \( \eta = 0 \),
\[
(W - b_H) < \phi(W - b_L)
\] (15)
and \( \phi_L > \frac{W - b_H}{W - b_L} \), the principal chooses \( b_L \) for any value of \( \phi \).

For very high numbers of conformists, e.g. for \( \eta = 1 \),
\[
(W - b_H) > (1 - 1)\phi(W - b_L)
\] (16)
or
\[
(W - b_H) > 0
\] (17)
the principal always pays \( b_H \).

In conclusion, if \( \phi_L > \frac{W - b_H}{W - b_L} \), the cut-off value under which the principal chooses \( b_L \) is
\[
\eta < 1 - \frac{W - b_H}{\phi(W - b_L)} =: \bar{\eta}_P(\phi)
\]
for \( \eta \in [\max\{0, \hat{\eta}(\phi_L)\}, \hat{\eta}(\phi_H)] \). There is a pooling equilibrium in which the principal pays \( b_L \) if \( \eta < \bar{\eta}_P(\phi_L) \). In case \( \eta > \bar{\eta}_P(\phi_L) \) and \( \phi_L > \frac{W - b_H}{W - b_L} \), the principal’s choices will result in hybrid equilibria.

\[\text{12 Only equilibria in pure strategies which satisfy the Cho-Kreps (1987) Intuitive Criterion are considered.}\]
5.1.2 Equilibria for optimistic prior expectation, $E[\phi] > \frac{1}{2}$:

When according to public information most steadfast agents are workaholics, i.e. for $E[\phi] > \frac{1}{2}$, paying $b_L$ has no signal value any more. The conformists will turn into workaholic agents in any case. This results in another equilibrium. When the sum of conformists and workaholics $\eta + \phi$ is sufficiently large, the principal should pay $b_L$ to save on the bonus payments. The critical value of $\eta(\phi)$ is such that

\[
\eta \theta(W-b_L) + (1-\eta)\phi \theta(W-b_L) > \eta \theta(W-b_H) + (1-\eta)(\phi \theta(W-b_H) + (1-\phi)\theta(W-b_H))
\]

or

\[
\eta \theta(W-b_L) + (1-\eta)\phi \theta(W-b_L) > \theta(W-b_H).
\]

Solving for $\eta$ yields the same cut-off value as in the initial case:

\[
\eta > \hat{\eta}_{OE}(\phi) = 1 - \frac{b_H - b_L}{(W-b_L)(1-\phi)}.
\]

When $\eta$ is above the value, the principal pays $b_L$. Else, she chooses $b_H$. The condition for this cut-off value to be greater than zero is

\[
\phi < \frac{W-b_H}{W-b_L}.
\]

When it is not fulfilled since $\phi_L > \frac{W-b_H}{W-b_L}$, the principal optimally pays $b_L$. Since $E[\phi] > \frac{1}{2}$, the conformists turn into workaholics anyway. Consequently, the cut-off value $\hat{\eta}_{OE}(\phi)$ provides the principal’s optimal policy even if $\hat{\eta}_{OE}(\phi_L)$ is smaller than zero.

Again, this cut-off value is also valid for very low values of $\eta$ since even for $\eta = 0$

\[
\phi_H \theta(W-b_H) + (1-\phi_H)\theta(W-b_H) < \phi_H \theta(W-b_L)
\]

\[
\phi_L \theta(W-b_H) + (1-\phi_L)\theta(W-b_H) > \phi_L \theta(W-b_L).
\]

For very high values of $\eta$ the situation is different than for $E[\phi] < \frac{1}{2}$, When $E[\phi] > \frac{1}{2}$, the conformists will turn into supporters in any case. Consequently, the principal can pay $b_L$ even for $\eta > \hat{\eta}_{OE}(\phi_L)$. For $\eta = 1$, she profits from paying $b_L$ since

\[
\Pi(b_H) = \theta(W-b_H) < \theta(W-b_L) = \Pi(b_L).
\]

6 Literature


