Motivation Systems and Pay Disclosure Policies: A Behavioral Approach

By Attila Lindner

Submitted to
Central European University
Department of Economics

In partial fulfilment of the requirements for the degree of Masters of Arts

Budapest, February 2010

Supervisor: Prof. Avner Ben-Ner
Assessor: Prof. Andrzej Baniak

Abstract

In this paper I analyze pay disclosure policies both from empirical and theoretical aspects. Even though information structure of the pay systems seems to be relevant, very few attempts have been made on understanding them. Combining Kőszegi (2006) with Ertac (2006), I develop a framework that endogenizes people's perception about themselves. These perceptions are influenced by pay disclosure policies through a social learning process. One important result is showing that higher cost induced by wage secrecy can be beneficial. On the other hand, the signals in the social learning process should be as precise as possible, indicating that wage secrecy also induce costs.

I would like to thank to my supervisor Avner Ben-Ner his invaluable comments and the devoted time on me and this paper. I am also indebted to Hedvig Horvath for the inspiring discussions, for proof reading and for helping me through the tedious calculations. Special thanks are due to my whole family and Marton Barta, Noemi Kreif, Zsolt Hajnalka for proof reading and valuable comments. All remaining errors belong solely to me.

Contents

Chapter 1

Introduction

The idea of "incentives" has always been a central notion of economics. Exploring the relevant factors that may influence people's behavior is one of the most important issues in understanding the world. Prior to the 1970's, economic research on incentive problems had mostly concentrated on price mechanisms. Since then, the emergence of theories on asymmetric information and incomplete contracts has deepened our understanding and raised new questions. One of the new and progressive research areas where these new models are widely used is personnel economics, a field in which rigorous economic analysis is applied to answer questions related to daily issues of human resource management (Lazear 1999; Lazear and Schaw 2007). Rapidly evolving, personnel economics has contributed a lot to our understanding of issues such as wage compression, performance-pay, turn-over, hiring practices, training and team-work (Lazear and Shaw 2007).

However, one particular question related to motivational systems has remained still poorly explored: pay disclosure policies. The wide-spread presence of wage secrecy policies on the one hand, and ignited debates in recent years on opening them up (both in academic (Bierman and Gely 2004; Edwards 2007) and in non-academic levels (e.g. Belkin 2008)) on the other, suggest that these policies are relevant elements of employees' motivational system. Moreover, the paucity of transparent wage policies makes the question even more interesting. In what circumstances is a transparent wage policy beneficial? What are the costs of these policies?[1]

This paper focuses on one particular aspect of optimal pay disclosure policies. The central goal of the analysis is to explain horizontal (non-)transparency (whether wages

[1] A good example for this is a case study of Harvard Business Review on this issue. The case is about a company where wages are revealed by accident which leads to social unrest within the firm. At the end of the study, four top managers are asked whether to make wages transparent at the long term. The managers disagree on the issue. In fact, the unconcealed goal of the case study is to incline MBA students to think about this controversial issue and figure out the main dilemmas (HBR 2001).

are known by colleagues at the same level), while vertical (non-)transparency (whether people are aware of their superiors' wages) will be left out of discussion. The separation is reasonable, since the disclosure policies in the two fields do not necessarily have the same roots. For instance, one could convincingly argue that observing peers' wages conveys relevant information about their productivity or relative chances of promotion, while this is not at all true for observing superiors' saleries.[2]

Most of the literature focuses on internal (within firm) explanations of pay disclosure policies. I also follow this tradition.[3] There are at least three explanations in the literature on why people's behavior may be influenced by disclosure of payment among peers: social preferences (or fair-wage effort hypothesis, e.g. Frank 1984, Akerlof and Yellen 1990), the tournament argument (Bronars 1986 cited by MacLaughin 1988), and the imperfect self-knowledge argument (Fang and Moscarini 2005; Ertac 2006). In Section 2, I evaluate the theoretical consistency of these theories and their relevance for understanding empirical evidence.

I have found that some of these explanations succeeded in explaining why wages are kept secret in most firms. However, all of these explanations have some difficulities, when we take into consideration that employees (managers) tend to overestimate their peers' wages (Lawler 1965; Milkovich and Anderson 1972; Mahoney and Weitzel 1978).[4] Existing theories predict that firms should use transparent wages if workers have this overestimation of wages. It is also interesting why agents accept lower than average perceived wages, while they think that they have higher than average abilities (Meyer 1975; Milkovich and Newman 1987, Kőszegi 2006).

The proposed explanation in this paper is closely related to the third explanation, i.e. the imperfect self-knowledge argument. Moreover, it is strongly based on Kőszegi's (2006) seminal paper. He introduces ego-utility that represents people's desire to have high perceived ability or talent. Kőszegi (2006) also endogenizes overconfident beliefs, he develops a framework where people have some scope for manipulating beliefs, and he proves that the average agent tends to have higher perceived ability than his real ability. The key mechanism is the follwing: agents with ego-utility (deriving utility directly from self-esteem) collect information about themselves in a biased way: they tend to stop gathering information when they have high perceived ability, and start

[2]However, it is not trivial whether open pay systems at the peer level can be introduced without vertical pay communication. I ignore this problem, since even if vertical and horizontal pay communication are difficult to separate, it is worth understanding whether wage secrecy is motivated by vertical or horizontal (or both) reasons. However, the literature often does not separate these two kinds of transparency that often leads to confusion.

[3]If the reader is interested in external, labor market explanations, see Danzinger and Katz (1997) and Bar-Isaac et al. (2008).

[4]Unfortunately, the little empirical research made on the perception of peers' wages was about managers. Thus, our knowledge how 'ordinary' employees perceive their peers' wages is limited.

seeking for information when they have low perceptions.

Apart from Kőszegi's work, another important idea used in my argumentation is Ertac's (2006) model on social learning. She argues that social learning is possible in circumstances where there are common shocks in workers' productivity. This indicates that agents can learn from observing each other's wages (or in her model, from signals).

Combining Kőszegi (2006) with Ertac (2006), I develop a framework that endogenizes people's perception about themselves. These perceptions are influenced by pay disclosure policies through the learning process. The model presented here has two methodological contributions. First, combining the two models mentioned above, I evaluate wage secrecy policies in terms of their costs and benefits. Second, I deviate from the standard principal-agent framework by dividing the total productivity of employees to (1) their own productivity (ability, talent) and (2) the productivity related to an entity external to them (firm, position, superiors or subordinates). The total output is the function of effort level multipled by total productivity and a white noise. I argue that total productivity disregarding a small shock can be easily observed (it is simply a ratio of output and effort), while the contribution of the two factors is much harder to distinguish.

While total productivity consists of two factors, agents derive ego-utility only from their perceptions of their own ability. This creates demand for belief manipulation: agents prefer situations where they have high perception of their own productivity even if it is an overestimation. The supply for belief manipulation is motivated by the signal collection problem introduced by Kőszegi (2006): agents can collect additional information about others' external productivity after a signal about their own productivity is observed. Since they are assumed to be risk lover at low level of abilities (and risk averse at high abilities) they tend to gather information if they have low level of perceived ability.

The source of the additional signal is other agents' total productivity. If agents can observe each other's effort, observing signals about total productivity can be interpreted as observing others' past wages, For social learning, it is required that workers' external productivity be exposed to common shocks (Ertac 2006). This social learning builds the bridge between wage transparency and Kőszegi's (2006) model on overconfidence.

The main finding of the model is that the principal introduces a postive cost of signal collection at the optimum. The intuition behind this result is quite simple: the higher cost makes the learning process more difficult indicating that only agents with bad news learn. This makes the distribution of beliefs skewed toward positive beliefs that gives higher pay-off to the principal. On the other hand, the result about the signal's preciseness is unintuitive: the model predicts that the optimal pay disclosure system applies signals as precise as possible. Therefore, the model has an ambigous prediction about wage secrecy. On the one hand, wage secrecy increases the cost of signal collection, which is beneficial for the principal. On the other hand, wage secrecy may make the signals imprecise and decrease the principal's profit.

Another interesting result is related to the link between agents' overconfidence and their perceptions of others' wages. It turns out that the correlation between agents' external productivities determines their opinions on others' wages. The same mechanism that drives overconfidence in abilities also leads to overconfidence in external productivity. In case of positive correlation between external productivities, this leads to an overestimation of others' external productivity and wages.

The paper proceeds as follows: Section 2 discusses some of the empirical evidence related to pay disclosure and the relevant theoretical literature. Section 3 builds up the model, while in Section 4 the main results are presented. Section 5 concludes.

Chapter 2

Related Literature

2.1 Empirical Evidence

First, I describe what wage transparency means. Baker et al. (1994) highlighted that companies often use administered wage policies that emphasize job assignment, level of education and age. This means that employers often determine a wage interval for each worker based on these obsevable factors without taking into consideration productivity differences within these groups. Day (2007) argues that in a survey of 1,000 WorldatWork member companies only a quarter released organization-wide salary ranges. If the survey is representative enough, this indicates that for most of the firms even administered wages are kept in secret. In addition to the administered pay, merit and performance-based reward structures are also used within firms (Baker et al. 1994). Most of the time, wage secrecy refers to this part of the pay.

The first empirical question is how prevalent wage secrecy rules are. Bierman and Gely (2004) argue that 30% of the private employers have specific rules prohibiting employees from discussing their pay with coworkers. On the other hand, one of every 14 employers has actively adopted pay "openness policy" and 51% of the employers reported not having any specific policy on pay secrecy or transparency. Another study of the 149 of Fortune 1,000 companies shows that only 3.5% had "open pay information system" (Lawler 2003 cited by Day 2007). Moreover, a survey of over 1,000 WorldatWork member companies found that only less than three percent made employee pay levels public (Scott et al. 2003 cited by Day 2007). We can conclude, based on this evidence, that wage secrecy is the prevailing form of pay communication systems.

The second empirical question is how wage secrecy affects employees' perceptions. Table 1 shows the results of studies looking at managers' opinions about others' wages. The results show the same pattern across the three studies: an average manager

TABLE 1.
Accuracy of respondent perceptions of compensation levels[a]

Respondent perceptions	Perceived < actual	Accurate (± 2 per cent)	Perceived > actual
Superior pay			
Lawler[b]	62.4%	10.3%	27.2%
Milkovich	56.0	4.0	40.0
Mahoney	45.8	3.9	50.3
Peer pay			
Lawler[b]	37.5	17.4	45.1
Milkovich	38.0	8.0	54.0
Mahoney	28.4	7.7	63.9
Subordinate pay			
Lawler[b]	28.8	11.5	59.7
Milkovich	16.0	5.0	79.0
Mahoney	30.9	5.8	63.4

[a]Perceptions were judged accurate if within ± $200 of actual compensation in the Lawler and the Milkovich and Anderson studies, and if within ± 2 per cent of actual in the present study.
[b]Lawler's data are based upon studies in three organizations and are representative of findings reported by him in other studies.
Source: Mahoney and Weitzel, (1978)

tends to overestimate his or her peers' salary. Based on these pieces of evidence, Lawler (1965) concludes that an open pay system would decrease workers' negative perception on their relative position and increase their self-esteem and productivity. Later studies claimed that Lawler (1965) may have reached the wrong conclusion. For instance, Mahoney and Weitzel (1978) and Milkovich and Anderson (1972) found that managers with the most accurate perception have the lowest satisfaction.

The results of Table 1 also suggests that people tend to be underconfident in terms of their relative ranking in wages. This latter is hard to interpret in light of direct evidence showing that managers (and employees) tend to be overconfident in their abilities (Meyer 1975; Milkovich and Newman 1987). These two pieces of evidence suggest that an average employee (manager) believes that he is in fact a very productive employee, while he is paid very badly. Two natural questions may be raised based on this observation: how can an agent be overconfident and underconfident at the same time and why do not employees with these perception move to another company, where they may expect fair-wage rewards? In this thesis I will focus on the first question, while Kőszegi (2006) alludes to the second.

The third area of empirical investigations is how to directly estimate the effect of pay disclosure policies on productivity and satisfaction. One of the few attempts is the study of Futrell and Jenkins's (1978) that examines the question on 508 pharmaceutical salesmen in a "before-after with control group" experimental design. They found that an open pay system increases salesman job performance and their satisfaction.

Another approach is to test pay disclosure policies in a laboratory environment, where the experimenter can control for all relevant factors. Charness and Kuhn (2004) found that subjects are unresponsive to co-workers wages and “employers”’ profit does not depend on the pay disclosure policy. Bambergen (2008) found a negative impact of pay secrecy on performance and profits.[1]

Both the empirical and laboratory results suggest that pay openness has positive impact on performance and firms’ profits. However, Futrell and Jenkins (1978) and Bambergen (2008) also discuss the limitations of their results. They highlight that a quite precise objective evaluation system was available in their experiments. The importance of the latter is also supported by Avolio’s and Manning’s (1985) study, which examined the impact of a sudden and unexpected pay disclosure in a university environment. They uniquely explore the heterogeneity in the reaction of different types of employees from support staff to university professors. They conclude that people with “high instrumentality perceptions attributed little importance to the actual disclosure of pay.” (Avolio and Manning’s (1985) p. 147)

This short summary highlights our lack of detailed knowledge on wage secrecy. We are far from understanding what characteristics lead to using open pay policies and what deters from them. Further research would highly contribute to our understanding of these issues.

2.2 Related Theories

In economic literature there are three main explanations focusing on internal factors of pay disclosure policies. The first refers to social preferences. The rapidly growing literature emphasizes how incentive systems and wage distributions are influenced by non-selfish preferences (e.g., Frank 1984; Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Fehr and Fischbacher 2002). Besides, and closely related to these ideas, some researches lay an emphasis on the role of workers’ perception about the fairness of their wages and effort level (Akerlof and Yellen 1990, Gan 2000, Charness and Kuhn 2004). Both of these approches claim that people’s non-selfish attitudes lead to wage compression within firm. However, wage equalization is costly for firms, since it also diminishes incentive motives of wages. This literature, especially Akerlof and Yellen (1990), Gan (2000) and Charness and Kuhn (2004), argues that wage secrecy

[1] However, it is rather puzzling that even if workers seemed to be reluctant against wage differences, employers responded to transparency with wage compression (Charness and Kuhn 2004). Hageman (2007) also found that open pay leads to wage compression at the horizontal level, but not at the vertical level.

eliminates (or mitigates) the effect of social comparisons or unfair perception through withholding relevant information. This can be beneficial, since the presence of social preferences does not indicate compressed wages. However, empirical evidence rejects that people with more information compare themselves to others more (Day 2007, Charness and Kuhn 2004).[2] Bambergen (2008) even argues that less information can make comparison even worse: people tend to believe that wages are secret because employers can conceal some unfairness by that. Evidence on managers' overestimation of co-workers' wages in a non-transparent environment, e.g. Lawler 1965; Mahoney and Weitzel 1972; Milkovich and Anderson 1978, underpin Bambergen's (2008) argument.

The second argument refers to tournament theory. Bronars (cited by MacLaughin 1988) argues that in dynamic tournaments intermediate information creates a "leader" and a "trailer" after the first stage. However, if these advantages are revealed by the employer, "trailer" feels low chance to win the tournament and therefore he decreases his effort. However, low level of effort on the "trailer" side indicates low level of effort on the"leader" side as well. This indicates that at later stages competition is less severe. Therefore, information revelation has negative consequence on incentives in a dynamic tournament environment. This theoretical explanation has strong testable implications. It suggests that in a tournament-type environment horizontal differences should be kept secret, while vertical differences should be transparent. It also suggests that firms in which internal labor market is important tend to have secret wages. In fact this is supported by empirical observations, such as that larger firms, where internal labor market is larger, are more likely to implement secret wage policies. On the other hand, many other observations are hard to explain based on this theory. For instance in many workplaces we cannot observe tournament-type situations (e.g. in case of unskilled workers), still wages are secret. Moreover, top executives – a typical example of tournament type workers (Lazear and Shaw 2007) – have transparent wages in most cases (Belkin 2008). Finally, and most importantly, managers' overestimation of their peers' wages also suggests that wage secrecy makes the situation even worse (Lawler 1965; Milkovich and Anderson 1972; Mahoney and Weitzel 1978).

The third type of explanations considers a standard principal agents framework and extends it with an assumption that agents have imperfect self-knowledge about their productivity. Fang and Moscarini (2005) analyze whether common wage policy or individual specific contracts are optimal. In their model, they assume that workers are not aware of their real productivity, but their principal knows it. If principal offers an agent specific contract, people can learn about their real productivity based on this contract. On the other hand, if principal offers the same contract to everybody,

[2] It is worth mentioning that the focus here is on horizontal wage secrecy. This argument may not be valid for vertical transparency.

no private information is revealed. The cost of this latter contract is the lack of individual specific incentive schemes. So the principal's decision is whether to offer an individual specific contract with information revelation and accurate incentives or a common contract without information revealation, but with inaccurate incentives. In a situation with unbiased beliefs the former is an optimal strategy, because an average agent can recieve accurate observations about his abilities while on average his beliefs about productivity are unchanged. On the other hand, if agents are overconfident, common wage schemes should be offered, since it is worth preserving agents' high beliefs, even if they do not have the right incentives. There are two deficiencies in this model. First, it is hard to understand, if the average worker is overconfident, why he still accepts a common wage scheme. Why do not these workers fight for individual specific contracts or just move to another firm, which offers them an individual specific contract? With making overconfidence endogenous, which means that overconfidence appears as a result of optimization like in Kőszegi (2006) or Bénabou and Tirole (2002), we can resolve that problem. Second, Fang and Moscarini (2005) did not analyze the potential role of wage secrecy, so they do not consider a third alternative that may dominate both common contracts and (open) individual contracts: secret indvidual specific contracts.

Another theoretical approach to disclosure policies of incentives schemes is that of Ertac's (2006), which offers a very elegant theoretical framework of social learning. She also assumes agents' imperfect self-knowledge, but in her model agents can learn about it with observing each other's output. The necessary condition for learning is the presence of common productivity shocks in employees' output. One example of these shocks is working on a similar task. If agents observe each other's output, they can learn more on the difficulty of a task. Knowing better the toughness of a task means also that an agent can learn more about his own ability, since he knows better that, for instance, his low output is the result of his low ability or of the difficulty of the task. Based on this model, Ertac (2006) derives that if agents (workers) are risk neutral, the principal should implement an open pay system, where social learning is possible. On the other hand, Ertac (2006) also derives that the principal may withhold information if agents are overconfident enough. The reason for this is that agents' overoptimistic beliefs increase their efforts and so the principal's profit. However, as long as social learning leads to more accurate perception of abilities, it decreases the level of overconfidence and therefore the profit as well. Both Ertac (2006) and Fang and Mascorani (2005) highlighted therefore the important relationship between pay disclosure policies and agents' overconfidence, while the problem with these approaches are the same: it is not really explained how and why this overconfidence emerges.

To resolve this problem, I build up a framework similar to Ertac (2006), but with endogenous overconfidence proposed by Kőszegi (2006), which model analyzes agents task choice with ego-utility. Ego-utility in his context means that agents derive direct utility from perceived ability. He develops a framework where agents can collect costless signals about their own abilities. Agents' preferences on ego-utility determines the way how they collect information: they tend to stop when they have high perceived ability, so the level of confidence endogenously emerges. Moreover, he also derives that an overconfident agent has less tendency to choose a challenging and more informative task, where the risk of unsuccessfulness is higher, since losing self-esteem (ego-utility) is very costly. On the other hand, some underconfident agents may participate in the challenging task, because they hope to be successful and gain ego-utility.

My present analysis uses only the signal collection part of Kőszegi's (2006) model and does not deal with task choice. However, the results related to task choice in Kőszegi (2006) are useful extension of the results presented here and help to resolve many deficiencies that previous explanations had.

This short survey of theoretical models highlighted the lack of coherent explanations of the widespread presence of wage transparent policies. In this paper, I plan to fill this gap and offer an explanation that incoporate endogeneous belief formation and social learning into a standard principal-agent framework. I follow Ertac (2006) and assume symmetric information between agent and principal (contrary to Fang and Moscarini 2005). This makes the analysis more tractable, since we do not have to deal with the strategic interactions related to the principal's information revelation. I extend Ertac (2006) model by endogenizing agents' overconfidence, strongly relying on Kőszegi (2006).

Chapter 3

The Model

In this chapter I develop a framework in which I analyze the relationship between overconfidence and pay disclosure policies.[1] I depart from the standard principal-agent framework, in three different ways. First, I separate agents' total productivity $(a+k)$ into ability (a) and circumstances induced productivity (k) (hereafter external productivity). The latter can be interpreted in many different ways. For instance if the agent is a middle-level manager, it can be the productivity of subordinates, the superiors' quality (managers at the higher level of the hiearchy) or the quality of capital (e.g. softwares) that supports the manager's work.

Second, I assume that agents have imperfect knowledge about their own abilities (a), about their external productivity (k) and about their total productivity $(a+k)$. I also assume that the contribution of the two factors to the total productivity cannot be separated from each other, neither by the principal, nor by the agent. Therefore, the principal must to reward agents based on total productivity. Third, I extend the agents' standard objective function with ego-utility proposed by Kőszegi (2006).

To be more specific, there is one principal and a continuum number of (ex-ante) identical agents with measure 1.[2] I assume that agents (and the principal) are risk neutral, therefore I can set aside the insurance motives of contracting.[3] I presume that the principal has no private information, so she has to offer the same contract to everybody. Agents do not interact strategically with each other. The only "interaction" between them is that they can collect information about each other's productivity.

The principal faces with the standard moral hazard problem: given agents' behav-

[1] The model presented here strongly relies on Kőszegi (2006).

[2] Agents are identical ex ante, before nature decides on their ability and extrenal productivity.

[3] This assumption is not unusual in the literature (e.g. Lazear and Rosen 1981, Holmström 1982, Ertac 2006, Kőszegi 2006). Moreover, Gibbons (1998) makes the following suggestion based on empirical evidence: "the tradeoff between incentives and insurance is (again) far from all that matters". It also turned out that the results presented in the next section do not depend on the risk neutrality assumption.

ior (incentive compatibility constraints and participation constraints) she maximizes her profit. However, in addition to the standard (linear) contracting instruments, she can also affect her profit through influencing agents' learning behavior by setting disclosure policies. To specify the problem in more details, the agents' problem is described first.

There are three stages in the agents' game. In stage 0, nature decides on all agents' underlying characteristics (denoted with $x_i = (a_i, k_i)$), such as ability (a_i), external productivity (k_i) and total productivity ($a_i + k_i$). Then all agents privately recieve a signal about their own total productivity ($a_i + k_i$). Agents update their beliefs based on these signals in a Bayesian way. In stage 1, agents decide whether to observe one of their randomly chosen collegue's signals (with an additional noise) or move on without aquiring any new information. After the additional signal is collected, beliefs are updated and agents move on. I will refer to this stage as the signal collection problem. In stage 2, agents choose their optimal effort based on their updated beliefs about ability and external productivity.

In stage 0, nature decides on random variables. These are agents' abilities (a_i) and external productivities (k_i). Abilities are normally distributed with $\overline{a}$ mean and σ_a^2, and they are independent of external productivities and other agents' abilities as well. External productivities are also normally distributed with $\overline{k}$ mean and σ_k^2 variance, however, there are correlation between them: $Corr(k_i, k_j) \neq 0 = \rho_{ij} = \rho_k$ if $i \neq j$ (where k_i and k_j are agent i's and agent j's external productivity, respectively). These correlations have an intuitive interpratation: positive correlation implies that all agents tend to be affected in a same way by external factors. One example for this is a bad superior (principal), who has some negative effect on all her subordinates (agents).[4] On the other hand, negative correlation implies that higher level of others' external productivity goes hand in hand with a lower level of the agent's external productivity. An example for this is managerial favoritism: a superior (principal) is in favor of some of her subordinates (agents). Assuming that the principal has a fix amount of resources to devide between agents, higher favoritism toward others goes hand in hand with lower level toward a particular agent.[5] While the sign of the correlation turns out not to be crucial for the principal, it has to be non-zero for

[4] The negative effect is not necessarily equally strong (correlation between external productivities is not equal with one) because of matching between employer and employees: some weakness of the superior may affect her subordinates differently.

[5] In this model there are inifinte number of agents, so the ex ante probablity distrubution of random characteristics is the same as the ex post distribution. However, this makes the argument presented here a little bit problematic, since information about one particular agent's productivity has infinitesimally small effect on the total resources that can be devoted to others. In fact, to deal with this problem, the measure of agents whose signal is observed should be handled carefully. However, I ignore this extension here, since it does not give any new qualitative result.

making social learning possible.

After nature draws a true pair of ability and external productivity for every agent, agents receive a signal $((a_i + k_i)^s)$ from normal distributions with mean of the true total productivity $(a_i + k_i)$ and with $\sigma^2_{\varepsilon_{(a+k)^s}}$ variance. One possible interpretation of this signal is that agents observe their past salary. The salary is determined by total output, which in a standard principal setting, is the function of total productivity, agents' effort and an unobservable shock. Therefore, apart from the unobservable noise (that is the error term of the signal), observation of the salaries conveys information about agents' total productivity $(a_i + k_i)$. Based on these signals, agents update their beliefs. I denote the random variable that represents agent i's perception after stage 0 by $\mathbf{X}_i^0 = (A_i^0, K_i^0)$, where A_i and K_i represents the perception about ability and external productivity, respectively. The distribution of the underlying characteristics and the signal structure ensure that $\mathbf{X}_i^0$ will be normal distributed (see lemma 1).

The essential part of the model is the signal collection problem in stage 1, where agents decide on whether to observe one of their peers' signal (that can be interpreted as his past salary) or stop information gathering and move on to the effort decision problem (stage 2). Asking about others' signals is costly. For instance, if it is prohibited to talk about wages, this cost can be huge as the agent may fear being caught. On the other hand, in a non-transparent environment it is just the opportunity cost of the time devoted to look up past salaries.

Agents decide on signal collection based on their expected utility. If an agent moves on without information gathering, his perception on underlying characteristics $(\mathbf{X}_i^0)$ remains the same, and based on these beliefs his expected utility in stage 2 is $E_{\mathbf{X}_i^0}\left[U^i\right]$. On the other hand, if he collects an additional signal, his belief about underlying characteristics may change due to the new information acquired. I denote the new perception of random variables with $\mathbf{X}_i^1 = (A_i^1, K_i^1)$. In Lemma 2 in the Appendix it is derived that $\mathbf{X}_i^1$ is also normally distributed. The change in beliefs may also change the expected utility obtained in stage 2: $E_{\mathbf{X}_i^1}\left[U^i\right]$. The agents therefore compare these two expected values based on initial perceptions $\mathbf{X}_i^0$. Formally, agents collect an additional signal if

$$E_{\mathbf{X}_i^0}\left[E_{\mathbf{X}_i^1}\left[U_i\right]\right] - c \geq E_{\mathbf{X}_i^0}\left[U_i\right],$$

where $\mathbf{X}_i^0$ represents agent i's perception about underlying characteristics after stage 0, $\mathbf{X}_i^1$ represents agent i's perception if he collects an additional signal, U^i is the utility derived from the optimal effort problem (see later) in stage 2, and c is the cost of collecting an additional signal.

Observing others' signal (or salary) is informative as a result of correlation between external productivities. This represents the social learning component of the model. Moreover, the choice between collecting an additional signal or moving on immediately endogenizes manipulation of beliefs and leads to overconfidence. For capturing differences in pay disclosure policies, I assume that there are additional (normally distributed) noises in observing others' wages, so the variance of the signal about others' wages ($\sigma^2_{\varepsilon_{(a+k)^{-s}}}$) is higher than that about the own ($\sigma^2_{\varepsilon_{(a+k)^s}}$). In a transparent environment this difference is very small (or zero), while in a pay-secret environment it can be huge. Another, closely related factor of pay disclosure policies is the cost (c) of the signal collection problem. The principal can set the two factors, $\sigma^2_{\varepsilon_{(a+k)^{-s}}}$ and c, in addition to the choice of the optimal linear contract. In the main results of the paper, I assume that the principal has complete freedom to select c and $\sigma^2_{\varepsilon_{(a+k)^{-s}}}$. However, in a more realistic environment there is a relationship between c and $\sigma^2_{\varepsilon_{(a+k)^{-s}}}$. For instance, the principal can increase the cost of signal collection through forbidding conversation about wages. That policy not only decreases the number of conversations, but conversations about wages become also less clear, indicating that the $\sigma^2_{\varepsilon_{(a+k)^{-s}}}$ also increases. In Chapter 4, I partially touch the possibility of this situation.

In stage 2, agents maximize their utility based on their updated perceptions. This problem departs from the standard principal-agent problem by the agents' non-standard objective function, especially by the presence of ego-utility and external productivity.

Following Kőszegi (2006), I assume that agents derive direct utility from self-perception (so called ego-utility). Agents' total utility is the weighted sum of ego- and standard utility, so separability is assumed. People have imperfect knowledge about their own ability, so ego-utility is derived from the perceptions about abilities, A_i^0 or A_i^1.

The functional form of the ego-utility is a key assumption that leads to overconfidence. I assume that ability (a) is normally distributed. Moreover, the signal structure of the model ensures that the perceptions (A_i^0 or A_i^1), independently of the signal history, are also normally distributed (see Lemma 1 and Lemma 2). However, this indicates that the level of ego-utility is uniquely identified by the mean and the variance of the distribution. To simplify the problem, I assume that the variance does not enter directly into the ego-utility, therefore utility is defined just by the perceived mean of ability (denoted by $\overline{a}_i^0$ and $\overline{a}_i^1$).[6] As Kőszegi (2006) suggested, I assume that ego-utility exhibits Kahneman and Tversky's (1979) gain-loss utility with some initial expectation about ability ($\overline{a}$) as a reference point. This assumption drives agents' information-seeking attitude at low level of (perceived) ability (where the function

[6] Assuming separability and a linear relationship between preciseness of knowledge and ego-utility ($J\left(E_{\mathbf{X}_i^s}(a_i), Var_{\mathbf{X}_i^s}(a_i)\right) = J^1(E_{\mathbf{X}_i^s}(a_i)) - \zeta Var_{\mathbf{X}_i^s}(a_i)$), does not change the main results.

is convex), and information avoidance at high level of (perceived) ability (where the function is concave), a key behavioral attitude that drives most of the results. I use a rather general functional form for ego-utility that satisfies these properties:

$$J(x,\overline{a}) = \left\{ \begin{array}{lll} -g\left(\overline{a}-x\right) & \text{if} & x \leq \overline{a} \\ g\left(x-\overline{a}\right) & \text{if} & x > \overline{a} \end{array} \right\},$$

where $\overline{a}$ is the average level of ability, x is the (perceived) mean of ability ($\overline{a}_i^0$ or $\overline{a}_i^1$).

Imposing some structure on $g\left(.\right)$ is required to solve the model. I sum these up in Assumption 1:

Assumption 1

1. $g\left(\cdot\right)$ is strictly increasing concave function (so $-g\left(\cdot\right)$ is strictly decreasing convex);
2. $g(\overline{a}) = 0$;
3. $E\left[g\left(X\right)\right]$ exists if X is normally distributed;
4. $g\left(X\right)$ is differentable and satisfies the Lipschitz condition.

Only the first is a substantial assumption, it represents the gain-loss type of the utility described above. The second is required to make $g\left(\cdot\right)$ continuous at 0. It can be interpreted as a normalization: people with average expectation does not feel any extra ego-utility. The third and forth are just technical assumptions: the third ensures that the expected value of $J(x,\overline{a})$ exists, while the forth guarantees the interchangability of differentiation and integration.

There is one important difference in the ego-utility defined above compared to the one proposed by Kahneman and Tversky (1979). Here $J(\cdot,\cdot)$ exhibits symmetry between gains and losses: a gain increases the utility by the same degree as a same amount of loss decreases it. On the other hand, Kahneman and Tversky (1979) based on experimental evidence suggest loss aversion, which means that people derive more disutility from a loss than the utility they get from the same amount of gain. Even though this latter assumption would fit better to experimental evidence, I ignore it on behalf of analytical tractability[7].

The second important deviation from the standard model is the presence of external productivity. I assume linear technology, mathematically:

[7] One example for the function that satisfies Assumption 1 is the following:

$$J(x,\overline{a}) = \left\{ \begin{array}{lll} \exp\left(h\left(x-\overline{a}\right)\right)-1 & \text{if} & x \leq \overline{a} \\ -\exp\left(-m\left(x-\overline{a}\right)\right)+1 & \text{if} & x > \overline{a} \end{array} \right\},$$

The advantage of this specification is that the solution is in closed form. However, the calculations are very tedious without specifying more parameter values.

$$q(a_i, k_i, e) = (a_i + k_i)e_i + \epsilon,$$

where ϵ is a white noise with zero mean and σ_ϵ^2 variance, $(a_i + k_i)$ is the agent's total productivity (the sum of own ability and external productivity) and e_i is the effort level. As in the standard moral hazard case, the effort is non-observable by the principal. I also assume that agents have standard cost functions of effort ($\frac{e_i^2}{2}$). These assumptions imply that the agents' third stage problem is to maximize the following expected utility:

$$E_{\mathbf{X}_i^j}[U_i] = \max_{e_i} \alpha J(E_{\mathbf{X}_i^j}[a_i], \overline{a}) + E_{\mathbf{X}_i^j}[w(q(a_i, k_i, e_i)] - \frac{e_i^2}{2}.$$

As I already noted, agents (and the principal) are risk neutral. This latter implies that the optimal linear contract offered by the principal depends on the output proportionally ($w(q(a_i, k_i, e_i)) = q(a_i, k_i, e_i) - B$), where B is determined by the principal based on agents' participation constraints.[8] For simplifying the calculations, I assume that agents' outside opportunities give them zero utility indicating that agents' expected utility (aquired before all signals are received) is zero at the optimal level of B:

$$0 = E\left[E_{\mathbf{X}_i^j}[U_i] - jc\right].$$

Moreover, as long as the measure of agents is one, the principals' profit is equal to B.

Based on these assumptions the principal's optimal pay disclosure policy (c and $\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}}$) is the following:

$$\pi = \max_{c, \sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}}} E\left[\alpha J(E_{\mathbf{X}_i^j}[a_i], \overline{a}) + E_{\mathbf{X}_i^j}[q(a_i, k_i, e_i^*)] - \frac{{e_i^*}^2}{2} - jc\right],$$

where $j \in \{0, 1\}$ denotes the number of additional signals collected and e_i^* is defined by agents maximization behavior in the second stage:

$$e_i^* = \arg\max_{e_i} \alpha J(E_{\mathbf{X}_i^j}[a_i], \overline{a}) + E_{\mathbf{X}_i^j}[q(a_i, k_i, e_i)] - B - \frac{e_i^2}{2}.$$

Recall that $\mathbf{X}_i^j$ is the vector of perceptions about random variables after observing j additional signals. j is the function of the cost of signal collection (c) and preciseness of the signals ($\sigma^2_{\varepsilon_{(a+k)^{-s}}}$), and it is determined at the first stage (signal collection problem) by the following equation:

[8] The optimal linear contract is more complicated to calculate here, since the moral hazard problem is extended with a learning process. However, the basic idea should work: as long as agents' and the principal's risk preferences are the same, it is no worth distorting agents' behavior.

$$j = \left\{ \begin{array}{ll} j = 1 & \text{if } E_{\mathbf{X}_i^0}\left[E_{\mathbf{X}_i^1}\left[U_i\right]\right] - c \geq E_{\mathbf{X}_i^0}\left[U_i\right] \\ j = 0 & \text{otherwise} \end{array} \right\}.$$

Before solving the model it is worth specifying what I exactly mean by the term of overconfidence in the next chapters. Benoit and Dubra (2007) highlighted that there are many concepts of overconfidence. In fact, I use two of them in the present analysis. I refer to "overconfidence I" as a situation where the median agent has higher than average perceptions. It is worth emphasizing that without information manipulation, median agents would have average perceptions, so "over" is not just a meaningless attribute. Moreover, this notion of overconfidence helps to compare the prediction of the model with empirical evidence on managerial perceptions such as the one saying that 90% percent of the managers believe that they are better than average (Meyer 1975).

The second notion of overconfidence compares the average level of preceptions about abilities to the average level of real abilities. Agents exhibit "overconfidence 2" if the former is bigger than the latter.

Chapter 4

Results

In this chapter I present the results; the reader can find the derivations in the Appendix. The first part of this chapter is devoted to solving the agents' problem, while the second part discusses the findings related to the optimal pay disclosure policies.

4.1 Agents' behavior

In the model agents have two decision points. First, based on their received signals in stage 0, they decide on whether to gather more information or not (the so called signal collection problem). Second, agents maximize their utility based on their beliefs about their total productivity (the so called effort decision problem). The decision on signal collection at stage 1 depends on the expected utility at stage 2, so I turn to analyze stage 2 first.

The optimal effort problem is just solving the effort maximization problem of agent i given their perceptions about random variables ($\mathbf{X}_i^j$):

$$E_{\mathbf{X}_i^j}\left[U_i\right] = \max_{e_i} \alpha J(E_{\mathbf{X}_i^j}\left[a_i\right], \overline{a}) + E_{\mathbf{X}_i^j}\left[(a_i + k_i)\, e_i + \epsilon)\right] - \frac{e_i^2}{2},$$

where $j \in \{1, 2\}$. Since $E_{\mathbf{X}_i^j}\left[a_i\right]$ does not depend on e_i and ϵ has zero mean, the previous problem can be simplified to the following:

$$e_i^* = \arg\max_{e_i} \overline{(a+k)}_i^j e_i - \frac{e_i^2}{2},$$

where $\overline{(a+k)}_i^j \equiv E_{\mathbf{X}_i^j}\left[(a_i + k_i)\right]$ by definition (see Lemma 1).

From this, the optimal effort of agent i is

$$e_i^* = \overline{(a+k)}_i^1.$$

This expression indicates that the effort level depends on the expected value of total productivity. Given agents' perceptions, it is also easy to calculate their expected utility in stage 2:

$$E_{\mathbf{X}_i^j}[U_i] = \alpha J(\overline{a}_i^j, \overline{a}) - B + \frac{1}{2}\left(\overline{(a+k)}_i^j\right)^2, \tag{4.1}$$

where $\overline{a}_i^j \equiv E_{\mathbf{X}_i^j}[a_i]$.

Equation (4.1) and the fact that $\overline{(a+k)}_i^j = \overline{a}_i^j + \overline{k}_i^j$ highlight that the higher (perceived) ability, ceteris paribus, increases agents' expected utility in two ways: first it increases his ego-utility, second it increases his (expected) reward on effort.

Using the calculated expected utility the signal collection problem can be solved. The condition for gathering new information is:

$$E_{\mathbf{X}_i^0}\left[E_{\mathbf{X}_i^1}[U_i]\right] - c \geq E_{\mathbf{X}_i^0}[U_i]$$

Proposition 1 summarizes one of the key equations of the model.

Proposition 1 *In the first stage, agents decide to collect one more signal if the following inequality holds:*

$$\alpha\left(E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)\right) + \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right] - c \geq 0. \tag{4.2}$$

Moreover $Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$ *does not depend on i (and the signal history).*

Proof. See Appendix. ■

Equation (4.2) highlights that the decision about collecting an additional signal is based on three important factors: the expected change of ego-utility, the expected profit increase and the cost of signal collection. The cost of collection has a straightforward effect: a cost increase leads to less signal collection. A more interesting result is that the expected reward ($\frac{1}{2}Var_{\mathbf{X}^0}\left[\overline{(a+k)}_i^j\right]$) is always postive. The intuition behind that is the following: collecting an additional signal introduces additional uncertainty before the signal is received. However, this uncertainty is favorable since agents' utility is a quadratic function of productivity.

The first part of equation (4.2) captures the idea that agents compare the expected ego-utility to its present level. This is the only part that depends on the past history of signals, so that part determines who collects an additional signal and who does not. The convexity-concavity assumptions stated in Assumption 1 determine agents' decision and the consequences of their beliefs are summarized in Proposition 2-5.[1]

[1]One way to see the importance of assumption 1, is to consider the case if $J(x, \overline{a})$ is linear in x. In that case $E_{\mathbf{X}_i^0}\left[J\left(\overline{a_i}^1, \overline{a}\right)\right] = J\left(E_{\mathbf{X}_i^0}\left[\overline{a_i}^1\right], \overline{a}\right) = J\left(\overline{a_i}^0, \overline{a}\right)$ so the ego utility part plays no role at all.

Proposition 2 *If α is high enough (so ego-utility is an important part of equation (4.2)) and assumption 1 holds, then more than half of the agents have higher than average perception about ability and about external productivity after the first stage. So they exhibit "overconfidence 1" about their own ability and about their external productivity.*

Moreover, if agents' external productivities are positively correlated ($\rho_k > 0$), then more than half of the agents believe that others have higher than average level of external productivity, while if the external productivities are negatively correlated ($\rho_k < 0$) then more than half of the agents believe that others have lower than average external productivity.

Proof. *See Appendix.* ■

The idea behind Proposition 2 is the same as in Kőszegi (2006). Agents with good perceptions about their abilities tend to avoid information, since they are risk averse. On the other hand, after agents have received bad news about their abilities, risk loving preferences are in action: agents seek more information to avoid bad self-evaluation. The gain-loss utility, therefore, drives agents to stop when they have positive self-esteem. Similarly to Kőszegi (2006), "overconfidence 1" emerged even though agents update their beliefs perfectly rationally. This kind of behavior is interpreted as self-esteem manipulation.

However, the second part of Proposition 2 is a new result. The introduction of external productivity offers an explaination for the observed discrepancy between managers' perceptions about others' wages and about their own overconfidence. As I already noted in Chapter 2, the existing explanations of wage secrecy fails to work in an environment where managers belive that they earn less than average. However, most evidence on managerial perception (e.g. Lawler (1965), Milkovich and Anderson (1972) and Mahoney and Weitzel (1978)) suggests that this condition is valid in most managerial positions. It is not understood either how the overconfidence in own ability can be harmonized with evidence on underconfidence in own rewards. In the framework presented here, we can resolve these discrepancies. In fact, it turned out that the correlation between external productivity factors are crucial to understand how perception about own wages and others' wages evolve.

There are two possibilities to discuss. In the first, the correlation between external productivities is negative. An example for this is when a superior favors some of the workers, so she gives them better access to external resources with good quality, or she just appreciates more some workers' output. Since agents exhibit "overconfidence 1" in external productivity for the same reason as they do about their own ability, the negative correlation implies that they tend to underestimate others' external productivity and wages. On the other hand, if external factors are positively correlated, the level of external productivity is overestimated, and so are others' wages.[2]

Analyzing the principal's optimal decision problem about the cost and the noisiness of the signal (so determining the optimal pay disclosure policy) is crucial to understand how agents react to the changes in these variables. Proposition 3 and 4 summarize the comperative static results that constrain the optimizing principle.

Proposition 3 *Comparative statics 1: the effect of a ceteris paribus change in the cost of signal collection (c)*

If α is high enough and assumption 1 holds, then the measure of agents who have higher than average perception about ability exhibits an inverted U-shape as a function of c with the maximum at $c = \frac{1}{2} Var_{\mathbf{X}_i^0} \left[\overline{(a+k)_i^1} \right]$.

Proof. *See Appendix.* ■

Proposition 4 *Comparative statics 2: the effect of a ceteris paribus change in the preciseness of the signal* ($\sigma^2_{(a+k)^{-s}}$)

If α is high enough, assumption 1 holds, and $c < \frac{1}{2} Var_{\mathbf{X}_i^0} \left[\overline{(a+k)_i^1} \right]$, *then the higher the variance of the signal is, the more agents belive that they are better than average.*

If α is high enough, assumption 1 holds and $c \geqslant \frac{1}{2} Var_{\mathbf{X}_i^0} \left[\overline{(a+k)_i^1} \right]$, *then the higher the variance of the signal is, the less agents believe that they are better than average.*

Proposition 2 highlights the inverted U-shape relationship between the cost of signal collection and the level of "overconfidence 1". Proposition 3 shows that the effect of preciseness depends on the cost of signal collection: for high costs, an increase in the preciseness of the signals decreases the level of "overconfidence 1", while for low costs the effect is exactly the opposite.

Up until now the main focus was on "overconfidence 1" and it was shown that agents' signal collection decision leads to an outcome where the median belief is higher than average. Another important question is whether "overconfidence 2" emerges, so whether the average level of perceptions is higher than the average level of real abilities. The answer is no, the result is summarized in Proposition 5.

[2] Even though Proposition 2 seems to offer an explanation for overestimating others' wages, it is far from complete. For instance, the model also predicts that agents tend to believe that their superior is fair toward them, and bosses tend to be talented. These are not really realistic predictions and are against the intuition of belief manipulation, where agents tend to collect information supporting that their total productivity is low because of external factors and not because of their own ability. A possible future research direction is to introduce this argument somehow.

Proposition 5 *If assumption 1 holds, agents' average perception is equal to their average ability $\overline{a}$.*

Proof. *See Appendix.* ■

Proposition 5 highlights the boundary of the model presented here. Even though information collection is selective, the Bayesian update ensures that all beliefs are rational. Therefore, the average level of the beliefs cannot be different from the average level of abilities, even though the distribution of beliefs becomes skewed.

4.2 The optimal pay disclosure policy

In this section I shortly analyze the principal's optimal pay disclosure policy.[3] First, I consider the benchmark case without ego-utility. Then I turn to a situation where agents have ego-utility.

In case of $\alpha = 0$, ego-utility plays no role. The principal sets the optimal level of cost and noisiness to maximze her profit. She solves the following problem:

$$\pi_{\alpha=0} = \max_{c,\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}}} E\left[\frac{1}{2}\left(\overline{(a+k)}_i^j\right)^2 - jc\right],$$

where $j = \{0,1\}$ is determined by the noisiness and the cost of signals (see equation (4.2)). Proposition 6 highlights that pay openness is the optimal disclosure policy in case of this objective function.

Proposition 6 *In case of $\alpha = 0$, the principal sets $\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}} = 0$ and $c = 0$.*

Proof. See Appendix. ■

To understand the intutition behind this result, observe that an additional signal increases the heterogeneity of beliefs, so collecting more signals leads to higher variance in $\overline{(a+k)}_i^j$. This heterogeneity is prefered, since the principal's profit is a convex function of agents' beliefs and so the higher the variance is, the higher the expected pay-off is.

On the other hand, in case of $\alpha > 0$, the presence of ego-utility may change this result. The principal's profit function is the following:

$$\pi_{\alpha>0} = \max_{c,\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}}} E\left[\alpha J(E_{\mathbf{X}_i^j}\left[a_i\right], \overline{a}) + \frac{1}{2}\left(\overline{(a+k)}_i^j\right)^2 - jc\right].$$

[3] All results presented here is approximate, since I do not deal with the adjustment related to the agents who do not work after the signals received. These are the agents who perceive negative level of total productivity. However, if both ability and external productivity is far from zero, this approximation is very good.

It is clear that if α is small enough, the optimal policy is dominated by agents' standard utility, so we get back Proposition 6. On the other hand, if α is high enough, the standard part of the profit can be abandoned. In Proposition 7 that case is considered.

Proposition 7 *If α is high enough (so the standard part of the utility has a small effect on the principal's profit), the optimal pay disclosure policy is $\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}} = \sigma^2_{\varepsilon_{(a_i+k_i)^{s}}}$ and $c = \frac{1}{2} Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$.*

Proof. See Appendix. ■

Proposition 7 is the key result of the paper. It highlights that the presence of ego-utility may lead to a situation where the principal introduces a cost to prevent some of the agents from signal collection. In fact, only agents who got some negative news will collect an additional signal, and so they have a chance to receive some positive news and update their beliefs. Another consequence of the proposition is that more precise signals are favorable, since, paradoxically, they leave more space for belief manipulation. A relatively precise signal means that new signals have a big effect on agent's beliefs and therefore, additional collection is more risky. However, the agents who collect new information are all risk loving, so they prefer that higher risk.

In a more realistic environment, the principal may not control all instruments of pay disclosure policies. If there is a relationship between the cost and noiseness of signals, the principal faces a trade-off. Prohibiting discussions at some level is beneficial, but also has costs, since it increases the impreciseness of signals. In that case, the optimal disclosure policy depends on the parameters of the model.

Chapter 5

Conclusion

In this paper I analyzed pay disclosure policies both from empirical and theoretical aspects. Even though information structure of the pay system seems to be relevant, very few attempts have been made to understand them. One goal of this paper is to take the first steps filling this gap. In the paper, I developed a model of social learning with endogenous overconfidence based on Kőszegi (2006) and Ertac (2006). I assumed that agents have imperfect information about their own abilities and they derive utility from the perception about them. I introduced a distinction between external and own productivity that turned out to be important to resolve some inconsistencies in the literature. One of these is to coherently derive how self-enhancing beliefs can be consistent with the overestimation of co-workers' wages. To the best of my knowledge, this is the first paper that could explain this seemingly contradictory piece of evidence.

The main result of the paper is that some cost preventing agents from social learning may be beneficial. Moroever, the best environment for belief manipulation is that of receiving signals as precise as possible. This creates a trade-off between the cost and preciseness of signals, indicating that the optimal pay disclosure policy depends on the relationship between these two factors.

The most obvious problem with the present version of the model is that the resulted overconfidence is very "weak". Empirical evidence suggests that "overconfidence 2", where the average level of perceived beliefs is higher than the real average, also emerges in many circumstances. An additional problem is that agents do not only exhibit "overconfidence 1" in abilities but also in external productivities. Even though this latter can explain why agents overestimate others' wages, leads to many, empirically probably false implications, such as the one saying that managers tend to admit their superiors talent and believe that their boss is in favor toward them. These problems suggest that Baysian update may be a too strong assumption and other endogenous models of overconfidence may lead to more realistic arguments.

This model is just a first step toward a more coherent analysis of optimal wage setting policies when agents have ego-utility (Kőszegi 2006, Bénabou and Tirole 2002) and face with information gathering decisions. This approach has the advantage of coherently dealing with agents' biased self-evaluation, a concept that seems to be widely accepted by human resource practitioners. Further understanding of the issue of pay and other reward disclosure policies is also important to improve the working environment within firms, universities and schools.

Chapter 6

References

Akerlof, George. and Yellen, Janet (1990). “The Fair-Wage Effort Hypothesis and Unemployment,” *Quarterly Journal of Economics*, 105(2): 255-284.

Avolio Bruce J. and Manning, Michael R. (1985). ”The Impact of Blatant Pay Disclosure in an University Environment”,*Research in Higher Education*, Vol. 23, No. 2

Baker, George and Gibbs, Michael and Holmstorm, Bengt (1994). “The Wage Policy of a Firm” *The Quarterly Journal of Economics*, Vol. 109, No. 4

Bambergen, Peter (2008). “Pay Secrecy and Individual Task Performance: The Mediating Effects of Fairness Perceptions and the Moderating Effects of Interpersonal Competitiveness, *manuscript*, Tilburg University

Bar-Isaac, Heski and Jewitt, Ian and Leaver Clare (2008).”Information and Human Capital Management”, *NYU Working Paper No.* 2451/26059

Belkin, Lisa (2008). "Psst! Your Salary Is Showing", New York Times, 19\Agust\2008

Bénabou Roland and Tirole, Jean (2002) ”Self Confidence and Personal Motivation”, *Quarterly Journal of Economics*, August 2002, 117(3), 871-915

Benoit, Jean-Pierre and Dubra, Juan (2007) "Overconfidence?", *SSRN Working Paper Series*

Bierman, Leonard and Gely, Rafael. (2004). "Love, Sex and Politics? Sure. Salary? No Way":Workplace Social Norms and the Law, *Journal of Labor and Employment Law*, 6, 120-156.

Bolton, Gary E. and Ockenfels, Axel (2000). “ERC: Theory of Equity, Reciprocity, and Competition”, *American Economic Review*. 90: 166-193.

Charness, Gary and Kuhn, Peter (2004). ” Do Co-Workers’ Wages Matter? Theory and Evidence on Wage Secrecy, Wage Compression and Effort”, *IZA Discussion Paper*, No. 1417

Danzinger, Leif and Katz, Eliakim (1997). “Wage Secrecy as Social Convention”

Economic Inquiry, Vol. 35, 59-69

Day, Nancy (2007). "A Investigation into Pay Communication: Is Ignorance Bliss?", *Personal Review*, Vol. 36, No. 5

Edwards, Matthew A. (2005). "The law and social norms of pay secrecy", *Berkley Journal of Employment and Labor Law*, 26, 41-63.

Ertac, Seda (2006). "Social Comparisons and Optimal Information Revelation: Theory and Experiments", P.hd. dissertation UCLA

Fang, Hanming and Moscarini, Giuseppe (2005). "Morale hazard", *Journal of Monetary Economics,* 52. 749–777

Fehr, Ernst and Fischbacher, Urs (2002). "Why Social Preferences Matter - The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives", *The Economic Journal*, Vol. 112, No. 478, Conference Papers, pp. C1-C33

Fehr, Ernst and Schmidt, Klaus M. (1999). "A Theory of Fairness, Competition, and Cooperation", *The Quarterly Journal of Economics*, 114: 817-868.

Frank, Robert (1984). "Interdependent Preferences and Competitive Wage Structure", *Rand Journal of Economics*, Vol. 15, No. 4

Futrell, Charles M. and Jenkis Omer C. (1978). "Pay Secrecy Versus Pciy Disclosure for Salesmen: A Longitudinal Study". *Journal of Marketing Research*, Vol. 15, pp. 214-9

Gan, Lie (2000), "The Uncertain Fair-Wage Effort Hypotheses and Wage Secrecy",

Gibbons, Robert (1998). "Incentives in Organization", *The Journal of Economic Perspectives*, Vol. 12, No. 4

Hagemann, Petra (2007). "On the Impact of Transparent Wage Distributions and Horizontal Inequity Aversion in an Experimental Labor Market", SSRN

HBR – Harvard Business Review Case Studies (2001). "When Salaries aren't Secret"

Holmström, Bengt (1982). "Moral Hazard in Teams", *The Bell Journal of Economics*, Vol. 13, No. 2, pp. 324-340

Kahneman, Daniel and Tversky, Amos (1979). "Prospect Theory: An Analysis of Decision under Risk", *Econometrica* 47, 263-291.

Kőszegi, Botond (2006). "Ego Utility, Overconfidence and Task Choice" *Journal of the European Economic Association*, 4(4), pp. 673-707.

Lazear, Edward P (1999). "Personnel Economics: Past Lessons and Future Directions. Presidential Address to the Society of Labor Economists", *Journal of Labor Economics*, Vol. 17, No. 2 (Apr., 1999), pp. iv-236

Lazear, Edward P and Shaw, Kathryn L. (2007). "Personnel Economics: The Economist's View of Human Resource", *NBER Working Paper* No. 13653

Lazear, Edward P and Rosen, Sherwin (1981). "Rank-Order Tournaments as Optimum Labor Contracts" Journal of Political Economy, 89:5, pp. 841-864

Lawler, Edwards E. (1965). ”Managers‘ Perceptions of their Subordinates‘ Pay and of their Superiors‘ Pay”, *Personnel Psychology*

Mahoney, Thomas. A. and Weitzel, William (1978). ”Secrecy and Managerial Compensation”, *Industrial Relations*, Vol. 17, No. 2

McLaughin, Kenneth J. (1988). ”Aspects of Tournament Models: A Survey”, *Research in Labor Economics*, Vol. 9, pp. 225-256

Meyer Herbert H. (1975). “The Pay-for-Performance Dilemma.” *Organizational Dynamics*,3(3), 39–65.

Milkovich, George T., and Newman, Jerry M. (1987). Compensation, 2nd ed. Business, Publications, Plano, TX

Milkovich, George T. and Anderson, Phillip H. (1972). ”Management compensation and secrecy policies”, *Personnel Psychology*, 25, 293-302.

Ruud, Paul A. (2000). "An Introduction to Classical Econometric Theory", New York: Oxford University Press.

Chapter 7

Appendix

Lemma 1

After observing a signal of $(a_i + k_i)^s$*, agent* i*'s perceptions of the underlying random variables* $X_i^0 = (\ A_i^0, K_i^0)$ *are normally distributed with the following means and variances:*

$$\overline{a}_i^0 \equiv E\left[a_i \right| (a_i + k_i)^s] = \overline{a} + \frac{\sigma_a^2}{\sigma_{(a+k)^s}^2}((a_i + k_i)^s - \overline{a+k}),$$

$$Var\left[a_i \right| (a_i + k_i)^s] = \sigma_a^2 - \sigma_a^2 \frac{\sigma_a^2}{\sigma_{(a+k)^s}^2},$$

$$\overline{k}_i^0 \equiv E\left[k_i \right| (a_i + k_i)^s] = \overline{k} + \frac{\sigma_k^2}{\sigma_{(a+k)^s}^2}((a_i + k_i)^s - \overline{a_i + k_i}),$$

$$Var\left[k_i \right| (a_i + k_i)^s] = \sigma_k^2 - \sigma_k^2 \frac{\sigma_k^2}{\sigma_{(a+k)^s}^2},$$

$$\overline{(a+k)}_i^0 \equiv E\left[a_i + k_i \right| (a_i + k_i)^s] = \overline{a+k} + \frac{\sigma_a^2 + \sigma_k^2}{\sigma_{(a+k)^s}^2}((a_i + k_i)^s - \overline{a_i + k_i}),$$

$$Var\left[(a_i + k_i)^s \mid (a_i + k_i)^s\right] = \sigma_{(a+k)^s}^2 - \left(\sigma_a^2 + \sigma_k^2\right) \frac{\left(\sigma_a^2 + \sigma_k^2\right)}{\sigma_{(a+k)^s}^2},$$

where $\sigma_{(a+k)^s}^2 \equiv \sigma_a^2 + \sigma_k^2 + \sigma_{\varepsilon_{(a+k)^s}}^2$.

Proof:

I use Ruud (2000): Let $\mathbf{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\sigma})$. If we partion $\mathbf{y} = \left[\begin{smallmatrix}\mathbf{y}_1 \\ \mathbf{y}_2\end{smallmatrix}\right], \boldsymbol{\mu} = \left[\begin{smallmatrix}\boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2\end{smallmatrix}\right]$ and $\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$, then $y_1 | y_2 \sim N(\mu_1 + \Omega_{12}\Omega_{22}^{-1}(\mathbf{y}_2 - \mu_2), \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{12}')$.

Using this theorem, perception of a can be calculated:

$\Omega_{11} = \left[\ \sigma_a^2\ \right], \Omega_{21} = \left[\sigma_a^2\right], \Omega_{12} = \left[\sigma_a^2\right],$

$\Omega_{22} = \left[\sigma_{(a+k)^s}^2\right].$

The inverse of Ω_{22} is

$\Omega_{22}^{-1} = \left[\frac{1}{\sigma^2_{(a+k)^s}} \right]$
Then using the exact form of $y_1|y_2$:

$$E\left[a_i|\left(a_i+k_i\right)^s\right] = \overline{a} + \frac{\sigma_a^2}{\sigma^2_{(a+k)^s}}((a_i+k_i)^s - \overline{a_i+k_i}),$$

$$Var\left[a_i|\left(a_i+k_i\right)^s\right] = \sigma_a^2 - \sigma_a^2\frac{\sigma_a^2}{\sigma^2_{(a+k)^s}}.$$

Using the same steps, it is easy to calculate the perceptions of $a_i + k_i$ and k_i as well.

Lemma 2

If agent i decides to observe an additional signal after observing his own one, his perceptions of the underlying random variables $X_i^1 = (\ A_i^1, K_i^1)$ are normally distributed with the following means and variances:

$$\begin{aligned}
\overline{a}_i^1 &\equiv E\left[a_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] = \\
&= \overline{a} + \frac{\sigma_a^2\sigma^2_{(a+k)^{-s}}}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}((a_i+k_i)^s - \overline{a+k}) - \\
&\quad -\rho_k\sigma_k^2\frac{\sigma_a^2}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}((a_i+k_i)^{-s} - \overline{a+k}),
\end{aligned}$$

$$Var\left[a_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] = \sigma_a^2 - \sigma_a^2\frac{\sigma_a^2\sigma^2_{(a+k)^{-s}}}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2},$$

$$\begin{aligned}
\overline{k}_i^1 &\equiv E\left[k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] = \\
&= \overline{k} + \frac{\sigma_k^2\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}((a_i+k_i)^s - \overline{a+k}) + \\
&\quad +\rho_k\sigma_k^2\frac{\left(\sigma^2_{(a+k)^s} - \sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}((a_i+k_i)^{-s} - \overline{a+k}),
\end{aligned}$$

$$\begin{aligned}
Var\left[k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] &= \sigma_k^2 - \sigma_k^2\frac{\sigma_k^2\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2} - \\
&\quad -\rho_k\sigma_k^2\frac{\left(\sigma^2_{(a+k)^s} - \sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2},
\end{aligned}$$

$$\overline{(a+k)}_i^1 \equiv E\left[a_i+k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] =$$
$$= \overline{a+k} + \frac{\sigma^2_{a+k}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}\left((a_i+k_i)^s - \overline{a+k}\right) +$$
$$+ \frac{\rho_k\sigma_k^2\left(\sigma^2_{(a+k)^s} - \sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}\left((a_i+k_i)^{-s} - \overline{a+k}\right),$$

$$Var\left[a_i+k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] = \sigma^2_{a+k} - \sigma^2_{a+k}\frac{\sigma^2_{a+k}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2} -$$
$$-\rho_k\sigma_k^2\frac{\left(\sigma^2_{(a+k)^s} - \sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2},$$

where $\sigma^2_{(a+k)^s} = \sigma_a^2 + \sigma_k^2 + \sigma^2_{\varepsilon_{(a+k)^s}}$, $\sigma^2_{(a+k)^{-s}} = \sigma_a^2 + \sigma_k^2 + \sigma^2_{\varepsilon_{(a+k)^{-s}}}$ *(and* $\sigma^2_{\varepsilon_{(a+k)^{-s}}} > \sigma^2_{\varepsilon_{(a+k)^s}}$*).*

Proof:

I again use Ruud (2000) here. First I calculate the perception of a:

$\Omega_{12} = \begin{bmatrix} \sigma_a^2 & 0 \end{bmatrix}$, $\Omega_{22} = \begin{bmatrix} \sigma^2_{(a+k)^s} & \rho_k\sigma_k^2 \\ \rho_k\sigma_k^2 & \sigma^2_{(a+k)^{-s}} \end{bmatrix}$ and

$\Omega_{22}^{-1} = \frac{1}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}\begin{bmatrix} \sigma^2_{(a+k)^{-s}} & -\rho_k\sigma_k^2 \\ -\rho_k\sigma_k^2 & \sigma^2_{(a+k)^s} \end{bmatrix}$.

Based on these matrices we can calculate the perceptions:

$$E\left[a_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] =$$

$$= \overline{a} + \frac{\sigma_a^2\sigma^2_{(a+k)^{-s}}}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}\left((a_i+k_i)^s - \overline{a+k}\right) -$$
$$-\rho_k\sigma_k^2\frac{\sigma_a^2}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}\left((a_i+k_i)^{-s} - \overline{a+k}\right)$$

$$Var\left[a_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] = \sigma_a^2 - \sigma_a^2\frac{\sigma_a^2\sigma^2_{(a+k)^{-s}}}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} - \left(\rho_k\sigma_k^2\right)^2}.$$

Using the same steps we can also calculate the perception of k:

$\Omega_{12} = \begin{bmatrix} \sigma_k^2 & \rho_k\sigma_k^2 \end{bmatrix}$, while Ω_{22} is the same as before. This implies that

$$E\left[k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right]=$$

$$\begin{aligned}
&= \overline{k}+\frac{\sigma_k^2\sigma^2_{(a+k)^{-s}}-\left(\rho_k\sigma_k^2\right)^2}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}}-\left(\rho_k\sigma_k^2\right)^2}((a_i+k_i)^s-\overline{a+k})+\\
&\quad+\frac{\rho_k\sigma_k^2\left(\sigma^2_{(a+k)^s}-\sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}}-\left(\rho_k\sigma_k^2\right)^2}((a_i+k_i)^{-s}-\overline{a+k}),
\end{aligned}$$

$$\begin{aligned}
Var\left[k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right] &= \sigma_k^2-\sigma_k^2\frac{\sigma_k^2\sigma^2_{(a+k)^{-s}}-\left(\rho_k\sigma_k^2\right)^2}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}}-\left(\rho_k\sigma_k^2\right)^2}-\\
&\quad-\rho_k\sigma_k^2\frac{\left(\sigma^2_{(a+k)^s}-\sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}}-\left(\rho_k\sigma_k^2\right)^2}.
\end{aligned}$$

In the same way it is also easy to calculate the perception of $a+k$.

Proposition 1

In the first stage, agents decide to collect one more signal if the following inequality holds:

$$\alpha\left(E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right]-J\left(\overline{a}_i^0,\overline{a}\right)\right)+\frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]-c\geq 0. \tag{4.2}$$

Moreover $Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$ *does not depends on* i *(and the past values of the signal)*

Proof:

Recall that agent i collects an additional signal if

$$E_{\mathbf{X}_i^0}\left[E_{\mathbf{X}_i^1}\left[U_i\right]\right]-c\geq E_{\mathbf{X}_i^0}\left[U_i\right].$$

At the optimum level of effort the expected utility is the following:

$$E_{\mathbf{X}_i^j}\left[U_i\right]=\alpha J(\overline{a}_i^j,\overline{a})+\frac{1}{2}\left(\overline{(a+k)}_i^j\right)^2-B.$$

Using this expression the following equation can be derived:

$$E_{\mathbf{X}_i^0}\left[\alpha J(\overline{a}_i^1,\overline{a})+\frac{1}{2}\left(\overline{(a+k)}_i^j\right)^2-B\right]-c\geq\alpha J(\overline{a}_i^0,\overline{a})+\frac{1}{2}\left(\overline{(a+k)}_i^j\right)^2-B,$$

which can be rearranged with simple algebra:

$$E_{\mathbf{X}_i^0}\left[\alpha J(\overline{a}_i^0,\overline{a})\right]+\frac{1}{2}E_{\mathbf{X}_i^0}\left[\left(\overline{(a+k)}_i^1\right)^2\right]-c\geq\alpha J(\overline{a}_i^0,\overline{a})+\frac{1}{2}\left(\overline{(a+k)}_i^0\right)^2 \tag{7.1}$$

(because of the properties of expectation and because B is constant).

First I calculate $E_{\mathbf{X}_i^0}\left[\left(\overline{(a+k)}_i^1\right)^2\right]$:

$$E_{\mathbf{X}_i^0}\left[\left(\overline{(a+k)}_i^1\right)^2\right] = Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right] + E_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]^2. \tag{7.2}$$

$E_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right] = \overline{(a+k)}_i^0$, since

$$\begin{aligned} E_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right] &\equiv E_{\mathbf{X}_i^0}\left[E\left[a_i+k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right]\right] \equiv \\ &\equiv E\left[E\left[a_i+k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right]|(a_i+k_i)^s\right] = \\ &= E\left[a_i+k_i|(a_i+k_i)^s\right] \equiv \\ &\equiv \overline{(a+k)}_i^0, \end{aligned}$$

where in the last but one equality I used the law of iterated expectations.

Now I turn to express $Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$:

$$\begin{aligned} Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right] &\equiv Var_{\mathbf{X}_i^0}\left[E\left[a_i+k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right]\right] \equiv \\ &\equiv Var\left[E\left[a_i+k_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right]|(a_i+k_i)^s\right] = \\ &= Var\left[\frac{\rho_k\sigma_k^2\left(\sigma^2_{(a+k)^s}-\sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}}+\left(\rho_k\sigma_k^2\right)^2}(a_i+k_i)^s\right] = \\ &= \left(\frac{\rho_k\sigma_k^2\left(\sigma^2_{(a+k)^s}-\sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}}+\left(\rho_k\sigma_k^2\right)^2}\right)^2\sigma^2_{(a+k)^{-s}}. \end{aligned}$$

where I used the results from Lemma 2, the properties of variance and the fact that in the second stage the signal $(a_i+k_i)^s$ is already known. Observe that $Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$ does not depend on any individual specific factor, since the variance of the signals are the same for everybody.

These results together with equation (7.2) implies that

$$E_{\mathbf{X}_i^0}\left[\left(\overline{(a+k)}_i^1\right)^2\right] = Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right] + \left(\overline{(a+k)}_i^0\right)^2.$$

Putting this result into equation (7.1) leads to the following inequality:

$$E_{\mathbf{X}_i^0}\left[\alpha J(\overline{a}_i^1,\overline{a})\right]+\frac{1}{2}\left(Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]+\left(\overline{(a+k)}_i^0\right)^2\right)-c \geq \alpha J(\overline{a}_i^0,\overline{a})+\frac{1}{2}\left(\overline{(a+k)}_i^0\right)^2.$$

Simple alegbra (and the properties of expectation) leads to the desired expression:

$$\alpha\left(E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^0,\overline{a}\right)\right]-J\left(\overline{a}_i^0,\overline{a}\right)\right)+\frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]-c \geq 0.$$

Lemma 3

Let $E_{A_i^0}\left[\overline{a}_i^1\right] = \overline{a}_i^0$ *and* $Var_{A_i^0}\left[\overline{a}_i^1\right] = \sigma^2$. *If assumption 1 holds, then* $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)$ *is the function of* $\overline{a}_i^0$ *and it has the following properties:*

1. $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)$ *is a continuous function of* $\overline{a}_i^0$.

2. $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) = 0$ *if the expected value of* A_i^0 *is* $\overline{a}$, *formally,* $\overline{a}_i^0 = \overline{a}$.

3. *Let* $\overline{a}_i^0 < \overline{a}$, $\overline{a}_i^1 \mid A_i^0 \sim N(\overline{a}_i^0, \sigma^2)$ *and* $\overline{a}_i^1 \mid \widetilde{A_i^0} \sim N(2\overline{a}_i^0 - \overline{a}, \sigma^2)$. *Then* $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) = -\left(E_{\widetilde{A_i^0}}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)\right)$. *So* $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)$ *is symmetric to the origin.*

4. *Let* $\overline{a}_i^1 \mid A_i^0 \sim N(\overline{a}_i^0, \sigma^2)$. *If* $\overline{a}_i^0 < \overline{a}$, *then* $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) > 0$ *and if* $\overline{a}_i^0 > \overline{a}$, *then* $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) < 0$.

5. *If* $\overline{a} > \overline{a}_i^0$, *then* $\frac{\Delta\left(E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)\right)}{\Delta\sigma} > 0$ *and if* $\overline{a}_i^0 > \overline{a}$, *then* $\frac{\Delta\left(E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)\right)}{\Delta\sigma} < 0$.

Proof:

1. $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)$ *is a continuous function of* $\overline{a}_i^0$.

By Assumption 1, part 4, $J(x, \overline{a})$ is continuous in x. $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right]$ is also continuous, since both $J\left(\overline{a}_i^1, \overline{a}\right)$ and the density function of A_i^0 (normally distributed by Lemma 1) are continuous. However, this implies the desired statement.

2. $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) = 0$ *if the expected value of* A_i^0 *is* $\overline{a}$, *formally,* $\overline{a}_i^0 = \overline{a}$.

Proof:

If $\overline{a}_i^0 = \overline{a}$, then $J\left(\overline{a}_i^0, \overline{a}\right) = g(\overline{a} - \overline{a}) = g(0) = 0$ by Assumption 1, part 2, so it is enough to show that $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] = 0$.

First observe that

$$\int_{-\infty}^{\overline{a}} g\left(\overline{a} - x\right)\varphi\left(x\right)dx = \int_{\overline{a}}^{\infty} g\left(x - \overline{a}\right)\varphi\left((2\overline{a} - x\right)dx,$$

where φ is the density function of the normal distribution with $\overline{a}_i^0$ mean and σ^2 variance. This can be easily shown by integration by substitution (with $h \equiv 2\overline{a} - x$):

$$\begin{aligned}
\int_{-\infty}^{\overline{a}} g\left(\overline{a}-x\right)\varphi\left(x\right)dx &= \int_{\infty}^{\overline{a}} g(h-\overline{a})\varphi\left(2\overline{a}-h\right)(-1)dh = \\
&= \int_{\overline{a}}^{\infty} g(h-\overline{a})\varphi\left(2\overline{a}-h\right)dh = \\
&= \int_{\overline{a}}^{\infty} g(h-\overline{a})\varphi\left(h\right)dh,
\end{aligned}$$

where in the last equality I used the fact that $\overline{a_i}^0 = \overline{a}$ and $h \geq \overline{a}$ imply that $\varphi\left(2\overline{a}-h\right) = \varphi\left(h\right)$.

Using this result, it easy to show that $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] = 0$:

$$\begin{aligned}
E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] &= \int_{-\infty}^{\infty} J\left(x, \overline{a}\right)\varphi\left(x\right)dx \equiv \\
&\equiv \int_{-\infty}^{\overline{a}} -g(\overline{a}-x)\varphi\left(x\right)dx + \int_{\overline{a}}^{\infty} g(x-\overline{a})\varphi\left(x\right)dx = \\
&= -\int_{\overline{a}}^{\infty} g(h-\overline{a})\varphi\left(h\right)dh + \int_{\overline{a}}^{\infty} g(x-\overline{a})\varphi\left(x\right)dx = 0.
\end{aligned}$$

Therefore,

$$E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) = 0 - 0 = 0.$$

3. *Let* $\overline{a}_i^0 < \overline{a}$, $\overline{a}_i^1 \mid A_i^0 \sim N(\overline{a}_i^0, \sigma^2)$ *and* $\overline{a}_i^1 \mid \widetilde{A_i^0} \sim N(2\overline{a}_i^0 - \overline{a}, \sigma^2)$. *Then* $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) = -\left(E_{\widetilde{A_i^0}}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)\right)$. *So* $E_{A_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)$ *is symmetric to the origin.*

Proof:

First I show that $J\left(\overline{a}_i^0, \overline{a}\right) = -J\left(2\overline{a} - \overline{a}_i^0, \overline{a}\right)$:

$$\begin{aligned}
J\left(\overline{a}_i^0, \overline{a}\right) &\equiv -g(\overline{a} - \overline{a}_i^0) = \\
&= g(2\overline{a} - \overline{a}_i^0) \quad \text{by definition of } g(\cdot) \\
&\equiv J\left(2\overline{a} - \overline{a}_i^0, \overline{a}\right).
\end{aligned}$$

I also show that $E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] = -E_{\widetilde{A_i^0}}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right]$:

$$E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] = \int_{-\infty}^{\infty} J\left(x,\overline{a}\right)\varphi\left(x\right)dx = \tag{L3.1}$$

$$= \int_{-\infty}^{\overline{a}} - g(\overline{a}-x)\varphi\left(x\right)dx + \int_{\overline{a}}^{\infty} g(x-\overline{a})\varphi\left(x\right)dx.$$

If $\overline{a_i}^0 < \overline{a}$ and $h \geq \overline{a}$ then $\varphi\left(2\overline{a}-h\right) = \varphi\left(h - 2\overline{a} + 2\overline{a_i}^0\right)$, so

$$\int_{-\infty}^{\overline{a}} g\left(\overline{a}-x\right)\varphi\left(x\right)dx = \int_{\overline{a}}^{\infty} g(h-\overline{a})\varphi\left(2\overline{a}-h\right)dh =$$

$$= \int_{\overline{a}}^{\infty} g(h-\overline{a})\varphi\left(h-2\overline{a}+2\overline{a_i}^0\right)dh.$$

Using these results, equation L3.1 can be rewritten as:

$$= \int_{\overline{a}}^{\infty} - g(h-\overline{a})\varphi\left(h-2\overline{a}+2\overline{a_i}^0\right)dh + \int_{-\infty}^{\overline{a}} g(\overline{a}-h)\varphi\left(h-2\overline{a}+2\overline{a_i}^0\right)dx =$$

$$= -\int_{-\infty}^{\infty} J\left(h,\overline{a}\right)\varphi\left(h-2\overline{a}+2\overline{a_i}^0\right)dh.$$

Moreover, observe that the density function of $\overline{a}_i^1 \mid \widetilde{A_i^0} \sim N(2\overline{a}_i^0 - \overline{a}, \sigma^2)$ is $\varphi\left(h-2\overline{a}+2\overline{a_i}^0\right) \equiv \psi(x)$, so

$$-E_{\widetilde{A_i^0}}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] \equiv -\int_{-\infty}^{\infty} J\left(h,\overline{a}\right)\psi(x)dh.$$

This result together with $J\left(\overline{a}_i^0,\overline{a}\right) = -J\left(2\overline{a}-\overline{a}_i^0,\overline{a}\right)$ imply the desired statement:

$$E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] - J\left(\overline{a}_i^0,\overline{a}\right) = -\left(E_{\widetilde{A_i^0}}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] - J\left(2\overline{a}-\overline{a}_i^1,\overline{a}\right)\right).$$

4. *Let* $\overline{a}_i^1 \mid A_i^0 \sim N(\overline{a}_i^0, \sigma^2)$. *If* $\overline{a}_i^0 < \overline{a}$, *then* $E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] - J\left(\overline{a}_i^0,\overline{a}\right) > 0$ *and if* $\overline{a}_i^0 > \overline{a}$, *then* $E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] - J\left(\overline{a}_i^0,\overline{a}\right) < 0$.

Proof:

This statement is the consequence of Jensen's inequality. Recall that Jensen's inequality states that

$$E(g(x)) > g(E(x))$$

if $g(\cdot)$ is strictly convex and

$$E(g(x)) < g(E(x))$$

if $g(\cdot)$ is strictly concave.

However, proving the statement is more complicated than to directly apply these inequalities. In the model, a part of the domain of function $J(\cdot,\cdot)$ is strictly concave, while the other part is strictly convex. However, it can be shown that if $\overline{a}_i^0 < \overline{a}$, then the convex part dominates, while if $\overline{a}_i^0 > \overline{a}$, the concave part dominates the direction of the inequality.

In this proof, I only focus on the case where $\overline{a}_i^0 < \overline{a}$. By the symmetry of $E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] - J\left(\overline{a}_i^0,\overline{a}\right)$ to the origin (see part 3 of this lemma), this also implies the statement for the other case, where $\overline{a}_i^0 > \overline{a}$.

By the definition of $J(\cdot,\cdot)$,

$$\begin{aligned} E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] &= P_{A_i^0}\left(A_i^0 \le \overline{a}\right) E_{A_i^0}\left[-g(\overline{a}-x)|A_i^0 \le \overline{a}\right] + \\ &+ \left(1 - P_{A_i^0}\left(A_i^0 \le \overline{a}\right)\right) E_{A_i^0}\left[g(x-\overline{a})|A_i^0 > \overline{a}\right]. \end{aligned}$$

Also by the definition of $J(\cdot,\cdot)$, if $\overline{a}_i^0 < \overline{a}$, then $J\left(\overline{a}_i^0,\overline{a}\right) = -g\left(\overline{a}-\overline{a}_i^0\right)$. Moreover, all expected values can be reformulated in a way that $E(X) = P(X \le h)E\left[X|X \le \overline{a}\right] + P(X > h)E\left[X|X > \overline{a}\right]$. Since $-g\left(\overline{a}_i^0,\overline{a}\right)$ is strictly convex

$$-g\left(\overline{a} - E_{A_i^0}\left[\overline{a}_i^1\right]\right) =$$

$$\begin{aligned} &= -g(\overline{a} - P_{A_i^0}\left(A_i^0 \le \overline{a}\right) E_{A_i^0}\left[\overline{a}_i^1|A_i^0 \le \overline{a}\right] - \\ &\quad - \left(1 - P_{A_i^0}\left(A_i^0 \le \overline{a}\right)\right) E_{A_i^0}\left[\overline{a}_i^1|A_i^0 > \overline{a}\right]) < \end{aligned}$$

$$\begin{aligned} < \; & -P_{A_i^0}\left(A_i^0 \le \overline{a}\right) g\left(\overline{a} - E_{A_i^0}\left[\overline{a}_i^1|A_i^0 \le \overline{a}\right]\right) - \\ & - \left(1 - P_{A_i^0}\left(A_i^0 \le \overline{a}\right)\right) g\left(\overline{a} - E_{A_i^0}\left[\overline{a}_i^1|A_i^0 > \overline{a}\right]\right). \end{aligned}$$

Substituting $J\left(\overline{a}_i^0,\overline{a}\right)$ with $-g\left(\overline{a} - E_{A_i^0}\left[\overline{a}_i^1\right]\right)$ leads to the following ineguality:

$$E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] - J\left(\overline{a}_i^0,\overline{a}\right) >$$

$$\begin{aligned} > \; & P_{A_i^0}\left(A_i^0 \le \overline{a}\right)\left(E_{A_i^0}\left[-g(\overline{a}-x)|A_i^0 \le \overline{a}\right] + g\left(\overline{a} - E_{A_i^0}\left[\overline{a}_i^1|A_i^0 \le \overline{a}\right]\right)\right) + \\ & \left(1 - P_{A_i^0}\left(A_i^0 \le \overline{a}\right)\right)\left(E_{A_i^0}\left[g(x-\overline{a})|A_i^0 > \overline{a}\right] + g\left(\overline{a} - E_{A_i^0}\left[\overline{a}_i^1|A_i^0 > \overline{a}\right]\right)\right). \end{aligned}$$

However, if both parts of the right handside are positive, then the left handside is bigger than zero, so the statement is proven.

The first part of the right handside is positive as an implication of Jensen's inequality:

$$E_{A_i^0}\left[-g(\overline{a}-x)|A_i^0 \leq \overline{a}\right] + g\left(\overline{a} - E_{A_i^0}\left[\overline{a}_i^1|A_i^0 > \overline{a}\right]\right) > 0,$$

where $-g(\cdot)$ is a strictly convex function. The second part of the right handside, $E_{A_i^0}\left[g(x-\overline{a})|A_i^0 > \overline{a}\right]+g\left(\overline{a} - E_{A_i^0}\left[\overline{a}_i^1|A_i^0 > \overline{a}\right]\right)$, is positive because the range of function $g(\cdot)$ is non-negative, and if x is non-zero, it is strictly positive.

5. *If* $\overline{a} > \overline{a}_i^0$, *then* $\frac{\Delta\left(E_{A_i^0}\left[J(\overline{a}_i^1,\overline{a})\right]-J(\overline{a}_i^0,\overline{a})\right)}{\Delta\sigma} > 0$ *and if* $\overline{a}_i^0 > \overline{a}$, *then* $\frac{\Delta\left(E_{A_i^0}\left[J(\overline{a}_i^1,\overline{a})\right]-J(\overline{a}_i^0,\overline{a})\right)}{\Delta\sigma} < 0.$

Proof:

I prove the case where $\overline{a}_i^0 < \overline{a}$. By the symmetry of $E_{A_i^0}\left[J\left(\overline{a}_i^1,\overline{a}\right)\right] - J\left(\overline{a}_i^0,\overline{a}\right)$ to the origin (see part 3 of this lemma), this also implies the statement for the other case, where $\overline{a}_i^0 > \overline{a}$.

Clearly, $J\left(\overline{a}_i^0,\overline{a}\right)$ does not depend on σ, so $\frac{\Delta J(\overline{a}_i^0,\overline{a})}{\Delta\sigma} = 0$. To show that $\frac{\Delta E_{A_i^0}\left[J(\overline{a}_i^1,\overline{a})\right]}{\Delta\sigma}$ is bigger than zero, first I argue that higher σ increases the distance from the mean for some (non-zero measure) "observations". To preserve symmetry, these observations need to be (almost surely) paired: one which is below the mean needs to have a pair which is above. Then I show that $J\left(\overline{a}_i^1,\overline{a}\right)$ is increased more by distance of the one above from the mean than it is decreased by distance of the one below from the mean. Therefore, the observations, whose distances are increased by the higher variance, increase the pay-off, so $\frac{\Delta E_{A_i^0}\left[J(\overline{a}_i^1,\overline{a})\right]}{\Delta\sigma} > 0$.

To show that a higher σ increases the distance from the mean, observe that $\overline{a}_i^1$ is symmetric to $\overline{a}_i^0$. This implies that the increasr of the standard deviation means that probability weight at the center (around the mean) falls while it increases at the tails. This can be only possible, if for some (non-zero measure, hypotetical) "observations"[1] the distance from the mean increases (and for all other observations it remains the same).

Let d be the distance from the mean. By the previous argument, if the variance increases for some (non-zero measure) observations the distance increases to $d' > d$. Moreover, to preserve symmetry, if $\overline{a}_i^0 - d$ becomes $\overline{a}_i^0 - d'$, then for almost all observations (that is, except for a zero measure set), there is another observation with $\overline{a}_i^0 + d$ becoming $\overline{a}_i^0 + d'$, otherwise symmetry would no longer hold.

The only thing that hass remained to be shown is that

[1] By hypotetical observations I mean here the observations that we would get if we drew a sample of continuum size. Since the sample size is continuum, it is also reasonable to say "positive measure of observations".

$$J\left(\overline{a}_i^0 + d, \overline{a}\right) + J\left(\overline{a}_i^0 - d, \overline{a}\right) < J\left(\overline{a}_i^0 + d', \overline{a}\right) + J\left(\overline{a}_i^0 - d', \overline{a}\right),$$

indicating that the utility acquired from the two pairs is bigger if the distance is bigger. There are three cases to consider, in the first two strict inequality holds, while in the third it holds with equality.

Case 1: $\overline{a}_i^0 + d < \overline{a}$ and $\overline{a}_i^0 + d' < \overline{a}$

In this case

$$J\left(\overline{a}_i^0 + d, \overline{a}\right) + J\left(\overline{a}_i^0 - d, \overline{a}\right) = -g(\overline{a} - \left(\overline{a}_i^0 + d\right)) - g(\overline{a} - \left(\overline{a}_i^0 - d\right))$$

and

$$J\left(\overline{a}_i^0 + d', \overline{a}\right) + J\left(\overline{a}_i^0 - d', \overline{a}\right) = -g(\overline{a} - \left(\overline{a}_i^0 + d'\right)) - g(\overline{a} - \left(\overline{a}_i^0 - d'\right)).$$

By the concavity of $g(\cdot)$, $g''(x) < 0$, so

$$\frac{g(\overline{a} - \left(\overline{a}_i^0 - d'\right)) - g(\overline{a} - \left(\overline{a}_i^0 - d\right))}{d' - d} < \frac{g(\overline{a} - (\overline{a}_i^0 + d)) - g(\overline{a} - \left(\overline{a}_i^0 + d'\right))}{d' - d}.$$

Rearranging the above inequality leads to the following inequality:

$$-g(\overline{a} - (\overline{a}_i^0 + d)) - g(\overline{a} - \left(\overline{a}_i^0 - d\right)) < -g(\overline{a} - \left(\overline{a}_i^0 + d'\right) - g(\overline{a} - \left(\overline{a}_i^0 - d'\right)),$$

which implies

$$J\left(\overline{a}_i^0 + d, \overline{a}\right) + J\left(\overline{a}_i^0 - d, \overline{a}\right) < J\left(\overline{a}_i^0 + d', \overline{a}\right) + J\left(\overline{a}_i^0 - d', \overline{a}\right).$$

Case 2: $\overline{a}_i^0 + d < \overline{a}$ and $\overline{a}_i^0 + d' > \overline{a}$

In this case

$$J\left(\overline{a}_i^0 + d, \overline{a}\right) + J\left(\overline{a}_i^0 - d, \overline{a}\right) = -g(\overline{a} - \left(\overline{a}_i^0 + d\right)) - g(\overline{a} - \left(\overline{a}_i^0 - d\right))$$

and

$$J\left(\overline{a}_i^0 + d', \overline{a}\right) + J\left(\overline{a}_i^0 - d', \overline{a}\right) = g(\left(\overline{a}_i^0 + d'\right) - \overline{a}) - g(\overline{a} - \left(\overline{a}_i^0 - d'\right)).$$

Let f be the distance between $\overline{a}_i^0 + d$ and $\overline{a}$. Then observe that

$$g(\overline{a} - \left(\overline{a}_i^0 - d'\right)) - g(\overline{a} - \left(\overline{a}_i^0 - d\right)) =$$

$$\begin{aligned} &= g(\overline{a} - \left(\overline{a}_i^0 - d'\right)) - g(\overline{a} - \left(\overline{a}_i^0 - (d+f)\right)) + \\ &\qquad + g(\overline{a} - \left(\overline{a}_i^0 - (d+f)\right)) - g(\overline{a} - \left(\overline{a}_i^0 - d\right)) \end{aligned}$$

and

$$g(\left(\overline{a}_i^0 + d'\right) - \overline{a}) - g(\overline{a} - \left(\overline{a}_i^0 + d\right)) =$$

$$= g(\left(\overline{a}_i^0 + d'\right) - \overline{a}) - g(\left(\overline{a}_i^0 + d'\right) - \overline{a} - (d' - (d + f)) + \\ + g(\left(\overline{a}_i^0 + d'\right) - \overline{a} - (d' - (d + f)) - g(\overline{a} - \left(\overline{a}_i^0 + d\right)).$$

Moreover, by the definition of f,

$$g(\left(\overline{a}_i^0 + d'\right) - \overline{a} - (d' - (d + f)) = g(0) = g(\overline{a} - \left(\overline{a}_i^0 + (d + f)\right)). \tag{L5.1}$$

By the concavity of $g(\cdot)$

$$\frac{g(\overline{a} - \left(\overline{a}_i^0 - (d + f)\right)) - g(\overline{a} - \left(\overline{a}_i^0 - d\right))}{f} < \\ < \frac{g(\overline{a} - \left(\overline{a}_i^0 + d\right)) - g(\overline{a} - \left(\overline{a}_i^0 + (d + f)\right))}{f}$$

and

$$\frac{g(\overline{a} - \left(\overline{a}_i^0 - d'\right)) - g(\overline{a} - \left(\overline{a}_i^0 - (d + f)\right))}{d' - (d + f)} < \\ < \frac{g(\left(\overline{a}_i^0 + d'\right) - \overline{a}) - g(\left(\overline{a}_i^0 + d'\right) - \overline{a} - (d' - (d + f))}{d' - (d + f)}.$$

The two equalities above imply

$$g(\overline{a} - \left(\overline{a}_i^0 - (d + f)\right)) - g(\overline{a} - \left(\overline{a}_i^0 - d\right)) + \\ + g(\overline{a} - \left(\overline{a}_i^0 - d'\right)) - g(\overline{a} - \left(\overline{a}_i^0 - (d + f)\right)) < \\ < g(\overline{a} - \left(\overline{a}_i^0 + d\right)) - g(\overline{a} - \left(\overline{a}_i^0 + (d + f)\right)) + \\ + g(\left(\overline{a}_i^0 + d'\right) - \overline{a}) - g(\left(\overline{a}_i^0 + d'\right) - \overline{a} - (d' - (d + f)).$$

Rearranging the inequality and using equation L5.1 lead to:

$$-g(\overline{a} - \left(\overline{a}_i^0 - d\right)) + g(\overline{a} - \left(\overline{a}_i^0 - d'\right)) < g(\overline{a} - \left(\overline{a}_i^0 + d\right)) - g(0) + g(\left(\overline{a}_i^0 + d'\right) - \overline{a}) - g(0).$$

Since $g(0) = 0$, this can be simplifed to the following inequality:

$$-g(\overline{a}-(\overline{a}_i^0-d))-g(\overline{a}-(\overline{a}_i^0+d))<g((\overline{a}_i^0+d')-\overline{a})-g(\overline{a}-(\overline{a}_i^0-d')),$$

which implies

$$J\left(\overline{a}_i^0+d,\overline{a}\right)+J\left(\overline{a}_i^0-d,\overline{a}\right)<J\left(\overline{a}_i^0+d',\overline{a}\right)+J\left(\overline{a}_i^0-d',\overline{a}\right).$$

Case 3: $\overline{a}_i^0+d\geq\overline{a}$ (and therefore $\overline{a}_i^0+d'\geq\overline{a}$)

In this case

$$J\left(\overline{a}_i^0+d,\overline{a}\right)+J\left(\overline{a}_i^0-d,\overline{a}\right)=-g(\overline{a}-(\overline{a}_i^0+d))-g(\overline{a}-(\overline{a}_i^0-d))$$

and

$$J\left(\overline{a}_i^0+d',\overline{a}\right)+J\left(\overline{a}_i^0-d',\overline{a}\right)=g((\overline{a}_i^0+d')-\overline{a})+g(\overline{a}-(\overline{a}_i^0-d')).$$

By assumption 1,

$$J\left(\overline{a}_i^0+d',\overline{a}\right)+J\left(\overline{a}_i^0-d',\overline{a}\right)=-g(\overline{a}-(\overline{a}_i^0+d'))-g(\overline{a}-(\overline{a}_i^0-d')).$$

However in Case 1, I showed that this implies

$$J\left(\overline{a}_i^0+d,\overline{a}\right)+J\left(\overline{a}_i^0-d,\overline{a}\right)<J\left(\overline{a}_i^0+d',\overline{a}\right)+J\left(\overline{a}_i^0-d',\overline{a}\right).$$

Proposition 2

If α is high enough (so ego-utility is an important part of equation (4.2)) and assumption 1 holds, then more than half of the agents have higher than average perception about ability and about external productivity after the first stage. So they exhibit "overconfidence 1" in their own ability and in their external productivity.

Moreover, if agents'external productivities are positively correlated ($\rho_k>0$), then more than half of the agents believe that others have higher than average level of external productivity, while if the external productivities are negatively correlated ($\rho_k<0$), then more than half of the agents believe that others have lower than average external productivity.

Proof:

First I prove that agents exhibit "overconfidence 1" in their own ability. There are two cases:

Case 1: $c\geq\frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$ (Recall that $\frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$ does not depend on individual specific factors, such as the true level of ability, the past realization of signals, and so it is not a function of $\overline{a}_i^0$.)

If $c \geq \frac{1}{2} Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, then by equation (4.2), none of the agents who have $E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) < 0$ collect any additional signal. Lemma 3, part 4 states that this holds for all agents with $\overline{a}_i^0 > \overline{a}$, so agents with higher than average perception (after the first signal is received) do not collect any additional signal. Since the number of agents is continuum, the distribution of $\overline{a}_i^0$ is almost surely the same as the underlying theoretical distribution. $\overline{a}_i^0$ is normally distributed with mean $\overline{a}$, so by the symmetry of the normal distribution, half (measure) of the agents have $\overline{a}_i^0 > \overline{a}$.

It is also known from Lemma 3, part 4 that $E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) > 0$ if $\overline{a}_i^0 < \overline{a}$. If α is high enough, then there exists $\overline{a}_i^{0\prime}$ such that

$$\alpha\left[E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^{0\prime}, \overline{a}\right)\right] > c - \frac{1}{2} Var_{\mathbf{X}_i^0}\left[\overline{(a_i+k_i)}^1\right] \tag{7.3}$$

for some agents. These agents choose to collect an additional signal. The only thing that we need to prove is that a positive measure of agents get so positive news that they update their beliefs from $\overline{a}_i^0 < \overline{a}$ to $\overline{a}_i^1 > \overline{a}$. This would imply that half of the agents (who received positive signal in the initial stage and do not collect) and a positive measure of agents (who collected a signal in the first stage and get so positive news that they update to $\overline{a}_i^1 > \overline{a}$) have higher than average beliefs. So "overconfidence 1" emerged (the measure of all agents is one).

To see that a positive measure of agents update their beliefs to $\overline{a}_i^1 > \overline{a}$, the results from Lemma 3 are used. First of all, by the continuity of $E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)$ (Lemma 3, part 1), there exists an $\varepsilon > 0$ such that equation (7.3) holds for all elements of $\left[\overline{a}_i^{0\prime} - \varepsilon, \overline{a}_i^{0\prime} + \varepsilon\right]$. By the continuum cardinality of agents and the normal distribution of $\overline{a}_i^0$, almost surely positive measure of agents are in the interval $\left[\overline{a}_i^{0\prime} - \varepsilon, \overline{a}_i^{0\prime} + \varepsilon\right]$. All of these agents collect an additional signal. By Lemma 2, for all level of beliefs, there exists a signal s^h, such that for all $s^h < (a_i + k_i)^{-s}$, agents update beliefs to $\overline{a}_i^1 > \overline{a}$. Since the signals are normally distributed, the probability of $s^h < (a_i + k_i)^{-s}$ is positive for all $a_i + k_i$. However, this implies that a positive measure of people who collect signals ($\overline{a}_i^0 \in \left[\overline{a}_i^{0\prime} - \varepsilon, \overline{a}_i^{0\prime} + \varepsilon\right]$) end up with $\overline{a_i}^1 > \overline{a}$.

Case 2: $c < \frac{1}{2} Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$.

The main idea behind this case is the same as before, but the derivation is more complicated. If $c < \frac{1}{2} Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, then by equation (4.2) agents who have $E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) > 0$ collect an additional signal. By Lemma 3, part 4, all agents with $\overline{a}_i^0 < \overline{a}$ collect an additional signal. However, if α is high enough, there exist an $\overline{a}_i^{0\prime\prime}$such that

$$\alpha\left[E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^{0\prime\prime}, \overline{a}\right)\right] < c - \frac{1}{2} Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]. \tag{7.4}$$

By continuity of $\alpha\left[E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)\right]$ (Lemma 3, part 1), there exist an

$\varepsilon > 0$ such that equation (7.4) holds for $\overline{a}_i^0 \in \left[\overline{a}_i^{0\prime} - \varepsilon, \overline{a}_i^{0\prime} + \varepsilon\right]$, so no one of them collect signals. By the continuum cardinality of agents and the normal distribution of $\overline{a}_i^0$, almost surely positive measure of agents are in that interval.

To prove that "overconfidence 1" emerges, I use the method of proof by contradiction. Assume that after signal collection "overconfidence 1" does not emerge. I showed that all agents with $\overline{a}_i^0 < \overline{a}$ collect signal and some positive measure of agents with $\overline{a}_i^0 > \overline{a}$ do not collect signals. Observe that if eveybody collected signal, then beliefs would be determined by Lemma 2 and they would be normally distributed with mean $\overline{a}$ ($\overline{a}_i^1$ is the linear function of normally distributed variables). However, if agents with $\overline{a}_i^0 > \overline{a}$ collect signals, there is a positive probability that they get sufficiently negative news and update their beliefs to $\overline{a}_i^1 < \overline{a}$. Since agents with $\overline{a}_i^0 > \overline{a}$ are continuum (positive measure), positive chance means that a positive measure of agents end up with $\overline{a}_i^1 < \overline{a}$.

By assumption, after agents solve the signal collection problem, "overconfidence 1" does not emerge. This means that at least half of the agents believe that they are worse than average. If all agents who do not collect collected, a positive measure of agents would join to the agents with lower than average beliefs. So agents' beliefs would exhibit "underconfidence 1". However, this is a contradiction, since if all agents collected signals, agents beliefs would have to be distributed symmetrically (normally) around $\overline{a}$.

Now, I turn to prove that agents' beliefs exhibit "overconfidence 1" in their external productivity after signal collection.

Case 1: $c \geq \frac{1}{2} Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$. As we saw in part 1 of this proof, only agents with negative signals in stage 0 collect signals. These agents have lower than average expectation about k as well (see. Lemma 1). The same argument as for a ensures that "overconfidence 1" emerges also in k.

Case 2. $c < \frac{1}{2} Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$.

Applying similar argument as for a, it can be shown that "overconfidence 1" emerges also in k.

The second part of the statement is the consequence of Bayesian update. If $\rho_k > 0$, agents who believe that their external productivity is higher than average, also believe that others' external productivities are higher than average (otherwise they would not update others' external productivity correctly). On the other hand, if $\rho_k < 0$, agents who think that their own external productivity is higher than average, believe that others' external productivity is lower than average. This (together with the fact that agents exhibit "overconfidence 1" in k) imply the desired statement.

Proposition 3

Comparative statics 1: the effect of a ceteris paribus change in the cost of signal *collection (c)*

If α is high enough and assumption 1 holds, then the measure of agents who have higher than average perception about ability exhibits an inverted U-shape as a function of c with the maximum at $c = \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$.

Proof:

I show that if $c < \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, the higher c is, the more agents believe that they are better than average. If $c < \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, then only agents with $\overline{a}_i^0 > \overline{a}$ never collect signals. Moreover, if c increased, more agents with $\overline{a}_i^0 < \overline{a}$ would collect signals indicating that almost surely more agents end up with $\overline{a}_i^1 > \overline{a}$, so the level of "overconfidence 1" increases.

On the other hand, if $c > \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, all agents with $\overline{a}_i^0 < \overline{a}$ collect signals. As c increases, the measure of agents who have $\overline{a}_i^0 > \overline{a}$ decreases. By the logic used in Proposition 2 Case 2, this implies that the level of "overconfidence 2" decreases.

By the continuity of $E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right)$, the previous two results imply that the maximum is achieved at $c = \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$.

Proposition 4

Comparative statics 2: the effect of a ceteris paribus change in the preciseness of the signal ($\sigma^2_{(a+k)^{-s}}$)

If α is high enough, assumption 1 holds, and $c < \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, then the higher the variance of the signal is, the more agents believe that they are better than average.

If α is high enough, assumption 1 holds and $c \geqslant \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, then the higher the variance of the signal is, the less agents believe that they are better than average.

Proof: The statement is the consequence of Lemma 3, part 5 and the idea used in Proposition 3.

If $c < \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, only agents with $\overline{a}_i^0 < \overline{a}$ collect signals. Changing $\sigma^2_{(a+k)^{-s}}$ decreases the variance of $\overline{(a+k)}_i^1$ and therefore, by Lemma 3, part 5, increases the number of people collecting an additional signal. The argument in Proposition 3 implies that the level of "overconfidence 1" increases.

If $c \geqslant \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$, all agents with $\overline{a}_i^0 < \overline{a}$ collect signals, while some of the agents with $\overline{a}_i^0 > \overline{a}$ do not. Changing $\sigma^2_{(a+k)^{-s}}$ decreases the variance of $\overline{(a+k)}_i^1$ and therefore, by Lemma 3, part 5, decreases the number of agents collecting an additional signal. The argument in Proposition 3 implies that the level of "overconfidence 1" decreases.

Proposition 5

If assumption 1 holds, agents' average perception is equal to their average ability $\overline{a}$.

Proof:

The statement is the consequence of Baysian update and the law of iterated expectations. Suppose that some of the agents collect signals, while others do not. Let μ be the measure of agents who collect signals, and let their average ability be $\overline{a}'$. Denote the average ability of those who do not collect with $\overline{a}''$. By the law of total expectation $\overline{a} = \mu\overline{a}' + (1-\mu)\overline{a}''$. Agents who collect signals update their beliefs as $E\left[a_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right]$, while others as $E\left[a_i|(a_i+k_i)^s\right]$. The average perception is the weighted expected value of these perceptions:

$$E\left[\mu E\left[a_i|(a_i+k_i)^s,(a_i+k_i)^{-s}\right]+(1-\mu)E\left[a_i|(a_i+k_i)^s\right]\right].$$

However, by the law of iterated expectations, this is exactly $\overline{a}$.

Proposition 6

In case of $\alpha = 0$*, the principal sets* $\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}} = 0$ *and* $c = 0$.

Proof:

If ego-utility plays no role, by Proposition 1, signal collection does not depend on individual specific factors. This implies that either everybody or nobody collects signals depending on $\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}}$ and c. The profit can be calculated by the following derivations.

By Lemma 1, $\overline{(a+k)}_i^0 = a + b(C-c)$, where $C \sim N\left(c,\sigma_c^2\right)$, $a = \overline{a}$ and $b = \frac{\sigma^2_{a+k}}{\sigma^2_{(a+k)^i}}$.

$$\begin{aligned} E\left[\left(\overline{(a+k)}_i^0\right)^2\right] &= \int [a+b(C-c)]^2\,dC = \int a^2 + b^2(C-c)^2 + 2ab(C-c)dC = \\ &= a^2\int dC + b^2\int (C-c)^2 dC + 2ab\int (C-c)dC = \\ &= a^2 + b^2\sigma_c^2. \end{aligned}$$

Plugging back a and b, the principal's profit if nobody collects signal can be calculated:

$$\begin{aligned} \pi_{\alpha=0} &\equiv \frac{1}{2}E\left[\left(\overline{(a+k)}_i^0\right)^2\right] \\ &= \frac{1}{2}\left[\overline{a}^2 + \left(\frac{\sigma^2_{a+k}}{\sigma^2_{(a+k)^i}}\right)^2\sigma^2_{(a+k)^i}\right] = \frac{1}{2}\left[a^2 + \frac{\left(\sigma^2_{a+k}\right)^2}{\sigma^2_{(a+k)^i}}\right]. \end{aligned}$$

By Lemma 2, $\overline{(a+k)}_i^1 = a + b_1(C-c) + b_2(K-k)$, where $C \sim N\left(c,\sigma_c^2\right)$, $K \sim N\left(k,\sigma_k^2\right)$, $a = \overline{a}$, $b_1 = \frac{\sigma^2_{a+k}\sigma^2_{(a+k)^{-i}} - \left(\rho_k\sigma_k^2\right)^2}{\sigma^2_{(a+k)^i}\sigma^2_{(a+k)^{-i}} - \left(\rho_k\sigma_k^2\right)^2}$, $b_2 = \frac{\rho_k\sigma_k^2\left(\sigma^2_{(a+k)^i} - \sigma_k^2\right)}{\sigma^2_{(a+k)^i}\sigma^2_{(a+k)^{-i}} - \left(\rho_k\sigma_k^2\right)^2}$.

$$
\begin{aligned}
E\left[\left(\overline{(a+k)}_i^1\right)^2\right] &= \int \left[a + b_1(C-c) + b_2(H-h)\right]^2 dC = \\
&= \int \left[a + b_1(C-c)\right]^2 + b_2^2(H-h)^2 + \\
&\quad +2\left[a + b_1(C-c)\right] b_2(H-h)dC \\
&= a^2 + b_1^2\sigma_c^2 + b_2^2\sigma_h^2 + 2Cov(a + b_1(C-c), b_2(H-h)) = \\
&= a^2 + b_1^2\sigma_c^2 + b_2^2\sigma_h^2 + 2b_1b_2Cov(C,H) = \\
&= a^2 + b_1^2\sigma_c^2 + b_2^2\sigma_h^2 + 2b_1b_2\rho_k\sigma_c\sigma_h \\
&= a^2 + (b_1\sigma_c + b_2\sigma_h)^2 - 2(1-\rho_k)\sigma_c\sigma_h.
\end{aligned}
$$

After plugging back a, b_1 and b_2, the principal's profit if everybody collects signal can be calculated:

$$
\begin{aligned}
\pi_{\alpha=0} &\equiv \frac{1}{2}E\left[\left(\overline{(a+k)}_i^1\right)^2\right] - c = \\
&= \frac{1}{2}\overline{a}^2 + \frac{1}{2}\left(\frac{\sigma_{a+k}^2\sigma_{(a+k)^{-i}}^2 - \left(\rho_k\sigma_k^2\right)^2}{\sigma_{(a+k)^i}^2\sigma_{(a+k)^{-i}}^2 - \left(\rho_k\sigma_k^2\right)^2}\right)^2 \sigma_{(a+k)^i}^2 + \\
&\quad + \frac{1}{2}\left(\frac{\rho_k\sigma_k^2\left(\sigma_{(a+k)^i}^2 - \sigma_k^2\right)}{\sigma_{(a+k)^i}^2\sigma_{(a+k)^{-i}}^2 - \left(\rho_k\sigma_k^2\right)^2}\right)^2 \sigma_{(a+k)^{-i}}^2 + \\
&\quad + \left(\frac{\sigma_{a+k}^2\sigma_{(a+k)^{-i}}^2 - \left(\rho_k\sigma_k^2\right)^2}{\sigma_{(a+k)^i}^2\sigma_{(a+k)^{-i}}^2 - \left(\rho_k\sigma_k^2\right)^2}\right)\left(\frac{\rho_k\sigma_k^2\left(\sigma_{(a+k)^i}^2 - \sigma_k^2\right)}{\sigma_{(a+k)^i}^2\sigma_{(a+k)^{-i}}^2 - \left(\rho_k\sigma_k^2\right)^2}\right)\rho_k\sigma_k^2
\end{aligned}
$$

From this latter equation (and from equation (4.2)), it is clear that if everybody collects an additional signal, the principal sets $c = 0$.

The only thing that has remained to be shown is that the principal earns higher profit if everybody collects an additional signal, so $E\left[\left(\overline{(a+k)}_i^1\right)^2\right] > E\left[\left(\overline{(a+k)}_i^0\right)^2\right]$. To see this, observe that

$$
\lim_{\sigma_{(a+k)^{-i}}^2 \to \infty} E\left[\left(\overline{(a+k)}_i^1\right)^2\right] = E\left[\left(\overline{(a+k)}_i^0\right)^2\right]
$$

and

$$
\frac{\partial E\left[\left(\overline{(a+k)}_i^1\right)^2\right]}{\partial \sigma_{(a+k)^{-i}}^2} < 0.
$$

which imply that $E\left[\left(\overline{(a+k)}_i^1\right)^2\right] > E\left[\left(\overline{(a+k)}_i^0\right)^2\right]$ for all $\sigma^2_{(a+k)^{-i}} < \infty$. The derivative of $E\left[\left(\overline{(a+k)}_i^1\right)^2\right]$ also implies that the principal chooses the lowest value of the preciseness of the signal: $\sigma^2_{(a+k)^{-i}} = \sigma^2_{(a+k)^{i}}$.

Proposition 7

If α is high enough (so the standard part of the utility has a small effect on the principal's profit), the optimal pay disclosure policy is $\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}} = \sigma^2_{\varepsilon_{(a_i+k_i)^{s}}}$ and $c = \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$.

Proof:

The principal maximizes $E\left[\alpha J(E_{\mathbf{X}_i^j}[a_i], \overline{a})\right]$. By Lemma 3, part 4, if $\overline{a}_i^0 < \overline{a}$, then $E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) > 0$, so agents' expected utility of signal collection is positive. The principal always prefers these agents to collect an additional signal. On the other hand, if $\overline{a}_i^0 > \overline{a}$, then $E_{\mathbf{X}_i^0}\left[J\left(\overline{a}_i^1, \overline{a}\right)\right] - J\left(\overline{a}_i^0, \overline{a}\right) < 0$, so the principal prefers these agents not to collect a signal. From Proposition 1, it is clear that $c = \frac{1}{2}Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$ ensures that only agents with $\overline{a}_i^0 < \overline{a}$ collect signals.

To show that $\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}} = \sigma^2_{\varepsilon_{(a_i+k_i)^{s}}}$, first observe that the lower $\sigma^2_{\varepsilon_{(a_i+k_i)^{-s}}}$ is, the higher $Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$ is. The formal proof is the following:

In the proof of Propoistion 1 it is derived that

$$Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right] = \left(\frac{\rho_k\sigma_k^2\left(\sigma^2_{(a_i+k_i)^s} - \sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} + \left(\rho_k\sigma_k^2\right)^2}\right)^2 \sigma^2_{(a+k)^{-s}}.$$

Calculating the derivatives with respect to $\sigma^2_{(a+k)^{-s}}$ leads to the following expression:

$$\begin{aligned}
\frac{\partial Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]}{\partial\sigma^2_{(a+k)^{-s}}} &= \left(\frac{\rho_k\sigma_k^2\left(\sigma^2_{(a+k)^s} - \sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} + \left(\rho_k\sigma_k^2\right)^2}\right)^2 - \\
&\quad -2\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}}\frac{\left(\rho_k\sigma_k^2\left(\sigma^2_{(a+k)^s} - \sigma_k^2\right)\right)^2}{\left(\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} + \left(\rho_k\sigma_k^2\right)^2\right)^3} \\
&= \left(\frac{\rho_k\sigma_k^2\left(\sigma^2_{(a+k)^s} - \sigma_k^2\right)}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} + \left(\rho_k\sigma_k^2\right)^2}\right)^2 \cdot \\
&\qquad \cdot\left(1 - \frac{2\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}}}{\sigma^2_{(a+k)^s}\sigma^2_{(a+k)^{-s}} + \left(\rho_k\sigma_k^2\right)^2}\right).
\end{aligned}$$

The first part of this expression is trivially positive. Moreover, the second part is negative since the absolute value of the correlation coefficient ($|\rho_k|$) is less than one and $\sigma_k^2 < \sigma_{(a+k)^s}^2 \leq \sigma_{(a+k)^{-s}}^2$. This leads to the desired statement:

$$\frac{\partial Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]}{\partial \sigma_{(a+k)^{-s}}^2} < 0.$$

At the optimal level of c, only agents with $\overline{a}_i^0 < \overline{a}$ collect signals. However, by Lemma 3, part 5, the expected utility of signal collection is higher if $Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]$ is higher. Since $\frac{\partial Var_{\mathbf{X}_i^0}\left[\overline{(a+k)}_i^1\right]}{\partial \sigma_{(a+k)^{-s}}^2} < 0$, the principal chooses as low $\sigma_{(a+k)^{-s}}^2$ as possible, so she sets the minimum value of $\sigma_{(a+k)^{-s}}^2$, that is $\sigma_{(a+k)^s}^2$.