

SCREENING OF POSSIBLY INCOMPETENT AGENTS AND WELFARE ANALYSIS WITHOUT COMMON PRIORS

NINA BARANCHUK* AND PHILIP H. DYBVIG[†]

November 08, 2011

Abstract

Accepting a contract with a high performance sensitivity is normally interpreted as a signal of high ability. However, an extremely high self-assessment may be an incompetent forecast by an incompetent worker. The authors do not want a portfolio manager who expects nearly riskfree returns of 50%/year! We study a model in which optimistic forecasters have low ability. In one type of equilibrium, a low performance sensitivity screens out the incompetent agents. In another, possibly not applicable for financial markets, agents are wealthy enough for the principal to select the incompetent agent who covers the downside (as in a vanity press). Doing welfare analysis of this phenomenon is challenging because agents have different priors and it is not clear what beliefs to use in evaluating welfare. Computing utility using agents' subjective beliefs, the outcome is relatively efficient because arbitraging the difference in beliefs improves welfare. However, in the objectivist view that there are true underlying probabilities, arbitraging beliefs adds noise and is inefficient. In either case, equilibrium tends to be second-best efficient since the agents' self-selection constraints, which use their own beliefs, undermine an objectivist social planner's efforts. Agents disagree with the social planner and are unwilling to move to the allocation in which the planner thinks they are better off. JEL Codes: G34 M51 M52

**Acknowledgments:* We are grateful to Anton Cheremukhin, Jun Liu, Chao Gu, Steve Ross, and seminar participants at Singapore Management University, University of Missouri-Columbia, National University of Singapore, and Southwest University of Finance and Economics. Philip H. Dybvig also thanks the Southwest University of Finance and Economics for support. Corresponding author. University of Texas at Dallas, P.O. BOX 830688 SM31, Richardson, TX 75083-0688. E-mail: nina.baranchuk@utdallas.edu.

[†]Washington University in St. Louis.

I Introduction

Models in which agents are selected (or self-select) based on ability typically assume that beliefs come from common priors.¹ In these models, performance sensitivity tends to attract agents of higher ability and screens out weaker candidates. For example, in the context of investment portfolio management, Bhattacharya and Pfleiderer (1985) show that, given common priors about managers' abilities, more competent managers self-select into positions with higher pay-performance sensitivity. In practice, however, there are no contests where portfolio managers are judged to be better if they take higher risk. We show that allowing agents to have different priors about their abilities permits a simple and plausible explanation for why this does not happen. Agents who are relatively pessimistic about their ability are likely to avoid jobs with any performance sensitivity, while agents who are very optimistic will self-select themselves for jobs with strong performance sensitivity. This may be a problem for employers expecting to attract good workers, especially when documented ability is scarce and self-assessed ability is great (as in the portfolio management industry). We build a model with these features, and we find two types of screening equilibria. If the manager's personal wealth is small compared to the amount of money at risk (as is typical of investment management), the equilibrium screens out the overly optimistic agents and gives the competent agent less performance sensitivity of compensation than would be optimal if it were not necessary to screen out the other agents. If, however, the manager has large personal wealth compared to what is at risk, employers may prefer to hire overly optimistic agents and require them to cover some of the losses. An agent with a strong belief in own ability and large wealth will be

¹See, for example, Spence (1973), Bhattacharya and Pfleiderer (1985), Koszegi and Li (2008).

A recent survey of screening and signalling models is provided in Riley (2001).

happy to guarantee the performance, and it will not matter if the employer thinks the performance will be poor. A good example of this is a vanity press that charges authors to publish their work.²

Our main model has a principal who seeks to hire an agent to fill a single position. We study equilibria in a pure screening setting (hidden exogenous type but no hidden effort) with two types of potential manager. From the perspective of the principal, one type is a good manager and is only slightly over-confident, while the other type is not a good manager but is extremely over-confident. Except possibly in welfare discussions, it does not matter who, if any one, has “correct” beliefs. We show that if the principal optimally wants to attract a more competent agent, the principal offers a wage schedule with little performance sensitivity, to make it unattractive to the incompetent but extremely confident agents. However, the principal may find attracting incompetent agents more profitable if the agents have large wealth (implying that the lower bound on wage is negative and large in absolute value). In this case, the principal offers a contract with a big embedded bet on output. Such a bet is attractive only to the extremely confident but incompetent type. We show that our results are robust to adding pessimistic and incompetent agents, and our results also extend to the setting with multiple positions.

The difference in beliefs makes welfare analysis in our model difficult and interesting. Welfare analysis is difficult because it is unclear what beliefs a social planner should be using to compute the expected utility that is an input to the welfare function or an assessment of Pareto optimality. We offer a general discussion of this issue but our analysis focuses on two particular approaches. One is a subjectivist approach that says

²Some authors who use a vanity press do not care about sales and simply want to produce a beautiful book, possibly just to give their friends. These are not the people in our models.

there are no objective probabilities and we should take the agents' subjective probabilities seriously and use them as inputs to the welfare function. Given the subjectivist approach, all equilibria are generally second-best efficient. The other approach we take is a pure objectivist approach that says that there are objective probabilities, in our analysis the same as the principal's probability assessments (which is consistent with assuming the principal is more sophisticated or more experienced than the agents). Given the objectivist approach, betting on differences in beliefs increases risk and possibly distorts investment, making it an inefficient way to transfer wealth to the agent with more accurate priors. Nevertheless, these bets are necessary for screening out the right agent types, and Pareto improvements are generally not available given the objectivist approach either.

It is possible to test our model empirically by studying the link between incentives in compensation and portfolio performance. As for most information models, we cannot observe all the variables of interest, and we must construct credible proxies. For example, while some studies find that stronger monetary incentives correspond to better performance (for example, Khorana, Serves, and Wedge (2007)), it is hard to tell from the evidence whether the stronger incentives proxy for effort or ability, and it is especially difficult to link the results to the managers' perceptions of their ability. A number of empirical measures of managerial overconfidence are offered in Malmendier and Tate (2005); evidence of managerial overconfidence is also reported in Ben-David, Graham, and Harvey (2007). However, they do not look at the link between their measures and CEO compensation packages, and that may be one place to look for evidence that can be used to test our model. In particular, one could test whether the overconfidence measures exhibit the positive relationship with observed pay-performance sensitivity implied by our model. It would also be interesting to see whether exogenous increases in pay-performance

sensitivity (driven, for example, by regulatory changes) affects the pool of applicants by attracting less talented but more overconfident applicants, especially in markets where overconfidence is a concern. To test the implications of our model, one could analyze whether (i) observed pay-performance sensitivity drops when screening out talented and less overconfident applicants becomes more important, and whether (ii) lower pay-for-performance indeed leads to an improvement in the quality of applicants. While it may be difficult to identify exogenous variation in the need to screen out talent and confidence, useful insights might be obtained if we make a reasonable assumption about when people will tend to be overconfident. For example, following an argument made in Malmendier, Tate, and Yan (2011), we can study an industry that has experienced a bubble period, when overconfidence and thus the need for screening is likely to be relatively large. Alternatively, following Malmendier and Nagel (2011), we can compare applicants who have lived through Great Depression to those who have experienced periods of high stock market returns.

Our assumption of heterogeneous prior beliefs represents a departure from the standard assumption of common priors. However, it seems closer to reality in many contexts, and, combined with Bayesian decision-making, it can produce numerous interesting and novel implications without abandoning the rational agent framework. For example, Van den Steen (2004) shows that agents with different priors tend to overestimate their ability to control the outcome, achieve success, and outperform others. Furthermore, differences in beliefs do not have to decrease over time, do not necessarily disappear in the limit even when agents are infinitely-lived, and are likely to persist when employment tenure is relatively short. Van den Steen (2005) focuses on team structure and argues that agents with similar priors are more likely to self-select into the same team. The effect of man-

agerial overconfidence on the firm investment policy and manager's welfare is studied in Gervais, Heaton, and Odean (2010). Adrian and Westerfield (2008) use a dynamic setting to analyze how disagreement between the agent and the principal impacts risk sharing. Dybvig, Farnsworth, and Carpenter (2010) note the possibility that agents with biased ability estimates may significantly affect the optimal compensation structure in a screening model, but do not develop this point. Landier and Thesmar (2009) use a screening model with overconfident borrowers to form predictions about debt maturity structure. Manove and Padilla (1999) use a model with overconfident borrowers to show that availability of collateral may destroy value because collateral helps stupid entrepreneurs fund bad projects.³ An argument for the widespread use of interest rate swaps based on differences in beliefs is presented in Arak, Estrella, Goodman, and Silver (1988). Hong and Stein (2007) argue that models with heterogeneous prior beliefs are particularly useful for improving our understanding of stock returns. This paper contributes to the literature by studying self-selection in the presence of fundamental disagreement about agents' abilities.

Section II describes the model and derives the equilibria. Section III discusses different definitions of social welfare absent common priors and shows how the different definitions change the welfare implications of our model. Section IV shows that our results are robust to various extensions, and section V concludes. The formal proofs are in the Appendix.

³Manove and Padilla (1999) may seem very similar to our paper because it is formulated as a signaling model. However, the information asymmetry is irrelevant in their equilibrium because, by design, the self-selection constraints are not binding, and all agents have the same contracts as they would get if the other types of agents were not there. Thus, their model does not have screening based on differences in prior beliefs, which is essential for our results.

II Model

There is a risk-neutral principal who wishes to hire one manager. There are two types of agents who can become managers: type o (optimistic, overconfident) and type c (conservative, competent). The type of each agent is known to the agent but not to the principal (unless revealed by the decision whether to apply for the position). Let π_o be the expected proportion of agents that is of type o , and let $\pi_c = 1 - \pi_o$ be the expected proportion of agents that is of type c . The principal is risk-neutral, while the agents are risk-averse with utility given by $\log(W)$, where W is the agent's wealth. The firm's output equals the manager's ability $a \in \{a_l, a_h\}$, where $a_h > a_l$.⁴ The agents' true ability is unknown to both the principal and the agents. Type o (resp. c) agents believe that their ability is high with probability q_o (resp. q_c), while the principal believes that their ability is high with probability f_o (resp. f_c). Beliefs about probabilities for various types are public knowledge (although agent types are private information), and we make no assumptions about the true distribution of the agents' abilities. For example, the equilibrium is the same whether type o agents have unrealistically optimistic expectations or the principal just doesn't appreciate how good they are. When we discuss the model results, however, we often take the view of the principal, as this view seems natural and helps the exposition.

We assume that $f_o < f_c < q_c < q_o$, which implies that both types of agents place a larger probability on their ability to be high than the principal does, and the principal believes that overly confident agents are less likely to have high ability. This is the interesting

⁴In a multi-stage model, we would probably want to set output equal to a function of ability and random noise, or otherwise everyone could learn ability from one draw on output. However, in our single-shot game, it does not matter much whether ability is uncertain but output equals ability, ability is certain but output given ability is uncertain, or ability is uncertain and output given ability is also uncertain.

case, although the intuition is similar when the competent agent is less optimistic than the principal, as we discuss in section IV. In that section, we also show that adding pessimistic ($q_p < q_c$) and incompetent ($f_p < f_o$) agents with a reasonably high reservation utility would not change the equilibrium because such agents would not accept the optimal contracts we describe, and the principal is happy about that. We additionally assume that agents have the same beliefs as the principal about all other agents' abilities. This assumption can be interpreted as saying that agents overestimate only their own ability, but not the ability of anyone else. In most of the paper, this assumption does not matter, but it helps the discussion of welfare in Section III by eliminating any demand for side bets between the principal and potential managers on other managers' success.

The principal advertises the managerial position with the wage schedule $w = (w_l, w_h)$, where compensation is w_l if the output is a_l and w_h if the output is a_h . We require the wage to be nondecreasing in the output: $w_h \geq w_l$, because a decreasing wage would give agents an incentive to destroy part of the output to receive a higher wage. All agents choose simultaneously whether to apply. If only one agent applies, the agent is hired. If several agents apply, the principal randomly picks an agent. We assume that, if all agents apply, the probability of hiring type o or c is equal to π_o and π_c respectively. If no agent applies, no one is hired and the principal receives a normalized zero payoff. We have in mind the solution concept of subgame perfect Bayesian Nash equilibrium in symmetric pure strategies, where by symmetric we mean agents of the same type play the same strategy. Since the assumed preferences have the single-crossing property, this implies the usual sort of screening equilibrium that can be analyzed graphically as in Figure 1, even with the differences in beliefs. Implicitly, we are ruling out schemes in which there is an application fee or compensation of unsuccessful applicants, on the grounds

that bookkeeping and processing costs would be prohibitive. We maintain these implicit restrictions for the social planner: this is not a paper about unemployment insurance.

We assume that each agent is endowed with the same initial wealth $W \geq 0$. All type o (resp. c) agents not hired by the principal work in the outside market and receive a wage with a certainty equivalent u_o (resp. u_c). Therefore, given other agents' application strategies, the agent of type $j \in \{o, c\}$ maximizes expectation of utility equal to

$$(1) \quad \begin{cases} q_j \log(W + w_h) + (1 - q_j) \log(W + w_l) & \text{if hired;} \\ \log(W + u_j) & \text{if not hired.} \end{cases}$$

The payoff for the principal from offering w is

$$(2) \quad \begin{cases} 0 & \text{if no agent applies} \\ f_o(a_h - w_h) + (1 - f_o)(a_l - w_l) & \text{if only type } o \text{ agents apply;} \\ f_c(a_h - w_h) + (1 - f_c)(a_l - w_l) & \text{if only type } c \text{ agents apply;} \\ \sum_{j=o,c} \pi_j (f_j(a_h - w_h) + (1 - f_j)(a_l - w_l)) & \text{if all agents apply.} \end{cases}$$

The following restrictions on the parameter space let us focus on cases that make our economic point.

Assumption 1. *a. The principal believes that all agents are less skilled than they think, and that more optimistic agents (type o) are less skilled than more conservative agents (type c):*

$$f_o < f_c < q_c < q_o.$$

b. Even a low-ability manager produces enough output to make it optimal to have a manager:

$$a_l > u_o.$$

c. Reservation utility of type o is not too large:

$$\frac{\log(W + u_o) - \log(W + u_c)}{q_o - q_c} < \log \left(\frac{(1 - f_c)q_c}{f_c(1 - q_c)} \right).$$

d. Reservation utility of type o is above that of type c :

$$u_o > u_c.$$

Assumption 1b is sufficient to ensure that the principal will hire someone in equilibrium. As we show in the proof of Theorem 1, assumption 1c is a necessary condition for ever having a pooling equilibrium where all agents apply. Given assumption 1c, assumption 1d becomes a necessary condition for ever hiring a type c agent in equilibrium. Assumptions 1c and 1d focus our attention on the economic points we want to make; for the curious reader, the end of this section briefly describes the solutions for the cases when assumptions 1c and 1d are violated.

The objective of the principal is to choose w that maximizes (2) subject to the individual rationality of agents' application decisions. Given our parameter restrictions specified in Assumption 1, there are three possible solutions to the principal's problem: (1) a solution with only type o agents applying; (2) a solution with only type c agents applying; and (3) a solution with all agents applying.⁵ The following theorem characterizes these solutions.

Theorem 1. *Given Assumption 1, any solution is of the type VP, TS, or NS defined below. The principal selects whichever type of solution has the highest payoff, for an overall payoff of $\max\{\Pi_{VP}, \Pi_{TS}, \Pi_{NS}\}$, where Π_{VP} , Π_{TS} , and Π_{NS} are defined below. All solution types obtain as we vary parameter values.*

VP (Vanity Press): *only the less competent (according to the principal) type o agents*

⁵Absent Assumption 1, there can be a solution in which no agent applies, for example if both agent types have reservation wages that are very high.

apply. The principal offers wage $w^* = (w_l^*, w_h^*)$ with

$$(3) \quad \begin{aligned} w_h^* &= (W + u_o) \left(\frac{f_o(1 - q_o)}{(1 - f_o)q_o} \right)^{q_o - 1} - W, \\ w_l^* &= (W + u_o) \left(\frac{f_o(1 - q_o)}{(1 - f_o)q_o} \right)^{q_o} - W; \end{aligned}$$

hires a type o agent and receives the following payoff:

$$(4) \quad \Pi_{vp} = f_o a_h + (1 - f_o) a_l - (W + u_o) \left(\frac{f_o}{q_o} \right)^{q_o} \left(\frac{1 - f_o}{1 - q_o} \right)^{1 - q_o} + W.$$

TS (Talent Screening). Only the more competent (according to the principal) type c agents apply. The principal offers wage $w^* = (w_l^*, w_h^*)$ with

$$(5) \quad \begin{aligned} w_h^* &= \exp \left\{ \frac{(1 - q_c) \log(W + u_o) - (1 - q_o) \log(W + u_c)}{q_o - q_c} \right\} - W, \\ w_l^* &= \exp \left\{ \frac{q_o \log(W + u_c) - q_c \log(W + u_o)}{q_o - q_c} \right\} - W; \end{aligned}$$

hires a type c agent, and receives the following payoff:

$$(6) \quad \begin{aligned} \Pi_{ts} &= f_c a_h + (1 - f_c) a_l - f_c \exp \left\{ \frac{(1 - q_c) \log(W + u_o) - (1 - q_o) \log(W + u_c)}{q_o - q_c} \right\} \\ &\quad - (1 - f_c) \exp \left\{ \frac{q_o \log(W + u_c) - q_c \log(W + u_o)}{q_o - q_c} \right\} + W. \end{aligned}$$

NS (No Screening). Both agent types apply. Let $f_m = f_o \pi_o + f_c \pi_c$ denote the principal's assessed probability that the agent hired will have high ability if both types apply. In this equilibrium, the principal offers wage $w^* = (w_l^*, w_h^*)$ with

$$(7) \quad \begin{aligned} w_h^* &= (W + u_c) \left(\frac{f_m(1 - q_c)}{(1 - f_m)q_c} \right)^{q_c - 1} - W, \\ w_l^* &= (W + u_c) \left(\frac{f_m(1 - q_c)}{(1 - f_m)q_c} \right)^{q_c} - W; \end{aligned}$$

hires a type c with probability π_c and hires type o with probability π_o and receives the following payoff:

$$(8) \quad \Pi_{ns} = f_m a_h + (1 - f_m) a_l - (W + u_c) \left(\frac{f_m}{q_c} \right)^{q_c} \left(\frac{1 - f_m}{1 - q_c} \right)^{1 - q_c} + W.$$

Proof: See the Appendix.

Corollary 1. *Let performance sensitivity of the manager's wealth be measured as follows:*

$$PPS = \frac{W + w_h}{W + w_l},$$

and let PPS_i denote the performance sensitivity given the wage in solution i . Then, $PPS_{VP} > PPS_{NS} > PPS_{TS}$.

Proof: From the expressions for wages in Theorem 1, performance sensitivity in NS, VP, and TS are given by, respectively,

$$\begin{aligned} PPS_{NS} &= \frac{(1 - f_m)q_c}{f_m(1 - q_c)}; \\ PPS_{VP} &= \frac{(1 - f_o)q_o}{f_o(1 - q_o)}; \\ PPS_{TS} &= \exp \left\{ \frac{\log(W + u_o) - \log(W + u_c)}{q_o - q_c} \right\}. \end{aligned}$$

The inequalities $PPS_{VP} > PPS_{NS} > PPS_{TS}$ follow from the above expressions after some algebra using Assumption 1.

The three possible solutions described in Theorem 1 are illustrated in Figure 1. In this figure, the two convex functions are the agents' indifference curves that leave them with their reservation utility levels. Because the optimistic agent's reservation utility is assumed to be above the more conservative agent's reservation utility, the lines cross above the 45 degree line. The straight lines in the figure represent the principal's isoprofit curves (the principal is risk-neutral). The lines have different slopes that depend on which types of agent the principal wishes to attract. The steepest slope corresponds to the case where the principal wishes to attract only the overconfident type o because the principal believes this type is less likely to succeed.

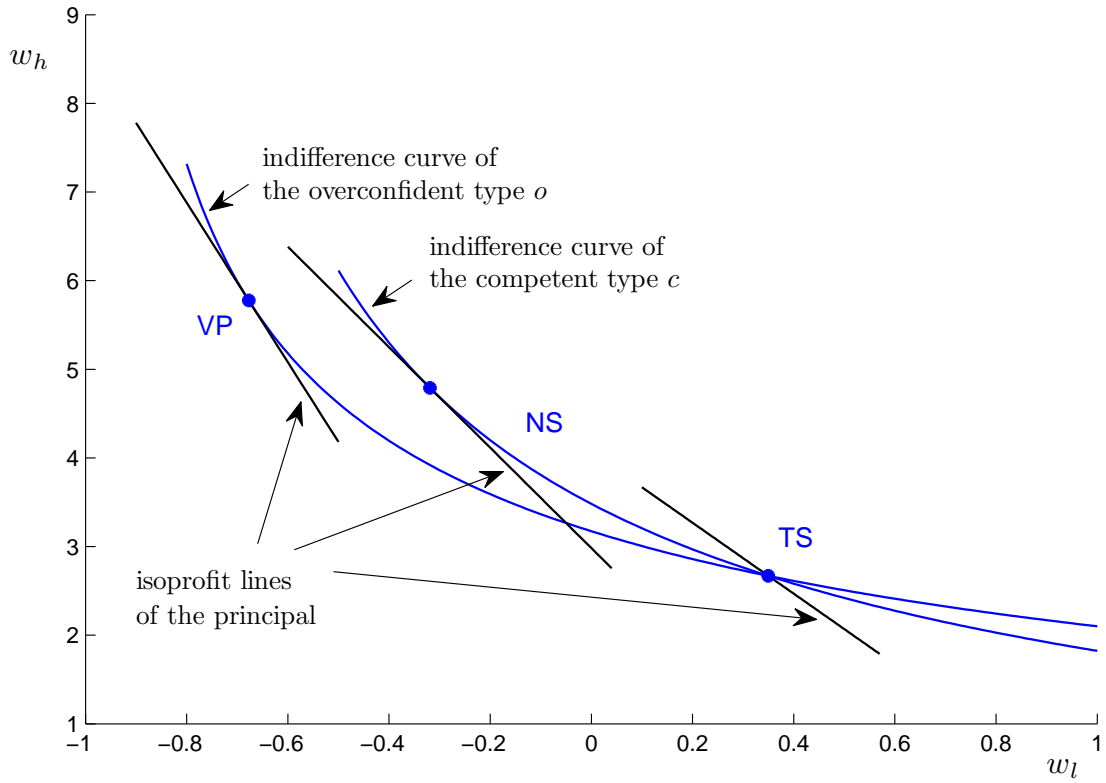


Figure 1: Model Analysis. The graph is built using the following parameter assumptions: $\pi_o = 0.5$, $\log(W + u_o) = 1$, $\log(W + u_c) = 0.9$, $q_o = 0.7$, $q_c = 0.6$, $f_o = 0.1$, $f_c = 0.2$, $W = 1$. The principal's isoprofit curves are straight lines whose slopes depend on which types of agent are expected to apply for the position.

The principal who has decided which types of agent to attract, wishes to do so using the smallest possible expected wage. The smallest expected compensation package that attracts type o agents is represented in the figure by the point called “VP”. This is the point where the principal’s isoprofit curve is tangent to the type o agents’ indifference curve. Similarly, the smallest expected compensation package that attracts type c agents is where the principal’s isoprofit curve is tangent to the type c agents’ indifference curve. This compensation package, however, would also attract type o agents (it is above the indifference curve that leaves type o agents with their reservation utility).

For the principal, the cheapest way to attract type c agents without attracting type o is to offer the compensation package denoted in the figure as “TS”. In this solution, even though type o is not hired, the presence of type o impacts the equilibrium compensation of the type c agent hired by the principal. For example, the equilibrium wage is more expensive to the principal when the type o agents are more optimistic (q_o is larger), provided these changes do not induce the principal to move from the TS equilibrium to another equilibrium (as they would do eventually). The presence of type o also lowers the performance sensitivity of the equilibrium wage; in the extreme case when the agents have the same reservation utility: $u_o = u_c$, the equilibrium wage is flat: $w_h = w_l = u_c$. In general, the performance sensitivity attracting the competent type c (PPS_{TS}) is below that attracting type o (PPS_{VP}), as stated in Corollary 1. This contrasts the results in Bhattacharya and Pfleiderer, where more able managers are attracted by higher performance sensitivity. In our model, better managers are more measured forecasters, and are therefore attracted by smaller performance sensitivity; the largest performance sensitivity occurs in the separating equilibrium attracting the overconfident type.

Selection of the manager in our model can be thought of as bundling two attributes:

productive efficiency and what bets are spanned. If the better type of manager (from the principal's perspective) is hired, then this means the claim on the other type of manager's output is not traded. This change in spanning (from spanning only a bad manager's output to spanning only a good manager's output) is important for the principal because the difference in beliefs is bigger for the bad manager. Thus, in equilibrium, the principal chooses to hire type o (Solution VP occurs) when the value of arbitraging the difference in beliefs with type o is large compared to the production efficiency loss. This tends to be true when the production efficiency loss is small, the difference in beliefs is large ($q_o \gg q_c$ is large), or initial wealth W is large, so that the agents' risk aversion is small.⁶ The principal, on the other hand, chooses to hire type c (Solution TS occurs) or to offer the pooling equilibrium wage (Solution NS occurs) when type c is very productive (f_c is large); the pooling wage is offered when there are too few type o agents (π_o is small) to justify screening.

Corollary 2. *While the principal always weakly prefers for the type c agent to have a lower reservation wage, lowering the reservation wage of the type o agent makes the principal strictly worse off in the TS region but strictly better off in the VP region.*

Proof. If the solution is of type VP ($\Pi_{VP} = \max\{\Pi_{VP}, \Pi_{TS}, \Pi_{NS}\}$), the principal's profit is given by (4), which is decreasing in u_o . If the solution is of type TS ($\Pi_{TS} = \max\{\Pi_{VP}, \Pi_{TS}, \Pi_{NS}\}$), the principal's profit is given by (6), which is increasing in u_o . \square

In our model, there is no hidden effort, which would make the model very messy to solve. However, the model has some features that make it qualitatively similar to what

⁶Formally, the absolute risk aversion at all level of net trade is smaller if initial wealth is larger.

we would have in a model with hidden effort. In particular, Assumption 1a implies that the principal and agent like some performance sensitivity no matter who is the manager, which should be similar to what would be optimal in a model with costly effort. However, it remains an open question exactly what the solution would look like in a model with both adverse selection and moral hazard.

When Assumption 1c is violated, the optimal wage that attracts type c lies below the point where the two indifference curves cross. In this case, there is no solution where both types apply, and if the principal chooses to hire a type c agent, the principal offers the wage specified in solution NS in Theorem 1, substituting f_c for f_m . When Assumption 1d is violated, the indifference curves for the two agent types cross below the 45-degree line, and thus the separating solution TS where a type c agent is hired becomes infeasible.

III Welfare Analysis

Traditional welfare analysis uses the agents' utility functions when computing social welfare or evaluating Pareto optimality. However, there are many circumstances in which it is unclear which utility function to use. For example, agents may be not fully aware of their preferences over unfamiliar outcomes, they may have time-inconsistent preferences, or they may have preferences over being free to make their own choices that go beyond preferences over outcomes. Under the commonplace assumption of von-Neuman-Morgenstern preferences and common prior beliefs, welfare is typically assessed either unconditionally using the common priors, or conditionally using posterior probabilities based on the common priors and pooled information. This approach requires modification when agents have different prior beliefs. In a pure subjectivist's view, there is no such thing as wrong

prior beliefs, and respecting each agent's own beliefs may make sense. On the other hand, in a pure objectivist's view, we may have objective views on what the true probabilities are, and may want to use them instead of agents' beliefs. Or, perhaps we should care some about subjective beliefs even in an objectivist world if part of happiness is based on anticipation and part on realizations (as in Dybvig and Rogers (2010)). In general, there are many approaches that take some agents' beliefs more seriously than others. In this paper, we focus on pure subjectivist and objectivist approaches for welfare analysis.

Presumably one point of doing welfare analysis is to think about policy issues, and therefore the appropriate approach to measuring welfare may depend on the context, our political philosophy, and our priors about whether well-intentioned government intervention will actually produce benefits net of costs. Most people would probably argue that, unless there is a reason to believe that agents are incompetent, we should respect their beliefs and use them in welfare calculations.⁷ If agents want to put at risk amounts of money they can afford to lose, it seems reasonable to let the agents make their own decisions.⁸ However, there are also circumstances where there is a reason to believe that agents may be acting against their own interest (for example, taking a large mortgage and failing to appreciate the likelihood of losing the house). It is not clear, in this case, to what extent one can assume that a third party (regulator) should interfere and act under the assumption of having a better understanding of the agent's beliefs. There is an argument that

⁷Some people would go further and argue that it is an inappropriate infringement of the right to freedom to impose the government's view on individual decisions. The definition and degree of incompetence is also a possible point of contention, and arguably everyone is incompetent to some degree.

⁸The view that this argument applies mostly for small potential losses can be justified by recognizing that agents may have some preference for freedom to determine their own fates. Of course, there are also other possibilities, for example, people who like not having to make a choice or people who care a lot about freedom to make big decisions even recognizing that they may be making big mistakes.

the government must intervene if an agent could make a big mistake, or if an agent fails to internalize substantial externalities causing damage to family and community. However, this is a dangerous argument that might be abused to justify almost any power grab by government.

In our setting, the main differences among these approaches to welfare analysis are in the value they attach to arbitraging differences in beliefs. Belief arbitrages are viewed as welfare-improving when the “subjectivist” welfare function respects the agents’ beliefs. On the other hand, “objectivist” welfare functions that use one common set of beliefs for all agents view belief arbitrage as a wealth transfer that may have negative side-effects. In our model, for example, we show that belief arbitrage may be viewed as a transfer of wealth from the hired agent to the principal, thus encouraging the principal to hire the less efficient worker. Moreover, belief arbitrage imposes additional risk on risk-averse agents, further reducing welfare.

Our model also offers recommendations for regulation. Under the reasonable assumption that the regulator cannot observe the agents’ types (the regulator does not know more than the principal), we show that regulation cannot lead to Pareto improvement. Although the equilibria in the model are not necessarily first-best Pareto optimal, they are second-best (as formally defined below) and improvement would require knowledge of the agents’ types; intervention by a regulator who is potentially less informed than the principal probably has only adverse or redistributive consequences.

In our formal welfare analysis, we take expectations of agents’ utilities prior to agents learning their types or employment status. We assume that all N agents are identical at the outset, $\pi_o N$ agents become type o , and the rest become type c (all agents share

these beliefs). Given that all agents are *a priori* identical, agent j expects to be hired with probability $1/N$. Let p_o (p_c) be the probability that a type o (c) agent is hired in a given allocation. Then, the probability that agent j becomes type o (c) and is hired equals p_o/N (p_c/N), while the probability that agent j becomes type o (c) and is not hired equals $\pi_o - p_o/N$ ($\pi_c - p_c/N$).⁹ Thus, the expected utility of agent j using the planner's beliefs (for evaluating the distribution of wages) can be written as:

$$\begin{aligned}
 U_j(w, p_o, p_c) &= \frac{p_c}{N} [q_c^w \log(W + w_h) + (1 - q_c^w) \log(W + w_l)] \\
 (9) \quad &+ \frac{p_o}{N} [q_o^w \log(W + w_h) + (1 - q_o^w) \log(W + w_l)] \\
 &+ \left(\pi_c - \frac{p_c}{N}\right) \log(W + u_c) + \left(\pi_o - \frac{p_o}{N}\right) \log(W + u_o).
 \end{aligned}$$

Expected utility (9) consists of four terms: the expected utility of the agent conditional on becoming a type c or type o and being hired, and the expected utility conditional on becoming a type c or type o and not being hired. In this expression, q_c^w and q_o^w represent the probabilities used by the social planner for type c and type o respectively. In particular, if the social planner takes the pure subjectivists' view, then the planner evaluates agents' expected utilities using their own beliefs, and we have $q_c^w = q_c$ and $q_o^w = q_o$. On the other hand, if the social planner takes the pure objectivists' view and evaluates agents' expected utilities using the principal's beliefs, we have $q_c^w = f_c$ and $q_o^w = f_o$ (taking the viewpoint of the principal is a natural choice, for example, if the principal is viewed as more sophisticated than the rest of the agents).

⁹Let t_j denote the type of agent j . Then, From Bayes formula,

$$\Pr(t_j = o, j \text{ is hired}) = \Pr(t_j = o | j \text{ is hired}) \Pr(j \text{ is hired}) = \left(\frac{[p_o/(\pi_o N)][\pi_o]}{1/N} \right) \left(\frac{1}{N} \right) = \frac{p_o}{N},$$

where the second equality is obtained using Bayes formula. Similarly,

$$\Pr(t_j = o, j \text{ is not hired}) = \left(\frac{[1 - p_o/(\pi_o N)][\pi_o]}{(N - 1)/N} \right) \left(\frac{N - 1}{N} \right) = (\pi_o - p_o/N).$$

The expected utility of the risk-neutral principal is equal to the expected profit, and thus can be written as follows:

$$(10) \quad \Pi(w, p_o, p_c) = [p_o f_o + p_c f_c](a_h - w_h) + (1 - [p_o f_o + p_c f_c])(a_l - w_l).$$

In our examples, we assume an objectivist planner has the same beliefs as the principal, so this expression is the same in both subjectivist and objectivist worlds.

There are a number of reasons why the equilibrium in this model is not first-best Pareto optimal. For example, an omniscient planner who knows agents types can provide unemployment insurance by giving a side payment to all agents who are not selected as the manager but are of the correct type. Moreover, since the type is viewed as random, there can be insurance against being a bad type. These things are unrelated to the point of our paper and therefore first-best Pareto optimality is not a useful benchmark. Still, before turning to second-best Pareto optimality, it is useful to examine a simple example to understand why the world looks less efficient in the first-best sense through the objectivist's prism. This example sidesteps the issues above by having a single manager of known type who can be hired by the principal.

Specifically, suppose there is only one agent whose utility and beliefs are those of type o in our model. Then, screening is not an issue, and the equilibrium wage is as described in Equilibrium VP . Figure 2 shows the subjective indifference curves of the agent and the principal using solid lines; the star marks the equilibrium wage. From the objectivists' point of view, the agent in this equilibrium is making a mistake by applying for the position instead of choosing the outside option, as it brings the agent expected utility below the reservation level (as evaluated by the social planner using the principal's beliefs). In Figure 2, this is captured by the dashed indifference curve, constructed using the

principal's beliefs to evaluate the agent's expected utility. This curve crosses the forty five degree line below the solid indifference curve. Because the agent is risk-averse while the principal is risk-neutral, the socially desirable wage would be independent of the firm's output. In particular, the objectivist social planner would prefer a flat wage, that leaves the principal's profit unchanged, to the equilibrium wage. This flat wage is represented in Figure 2 by w^{SP} , and is located at the intersection of the principal's indifference curve with the forty five degree line.

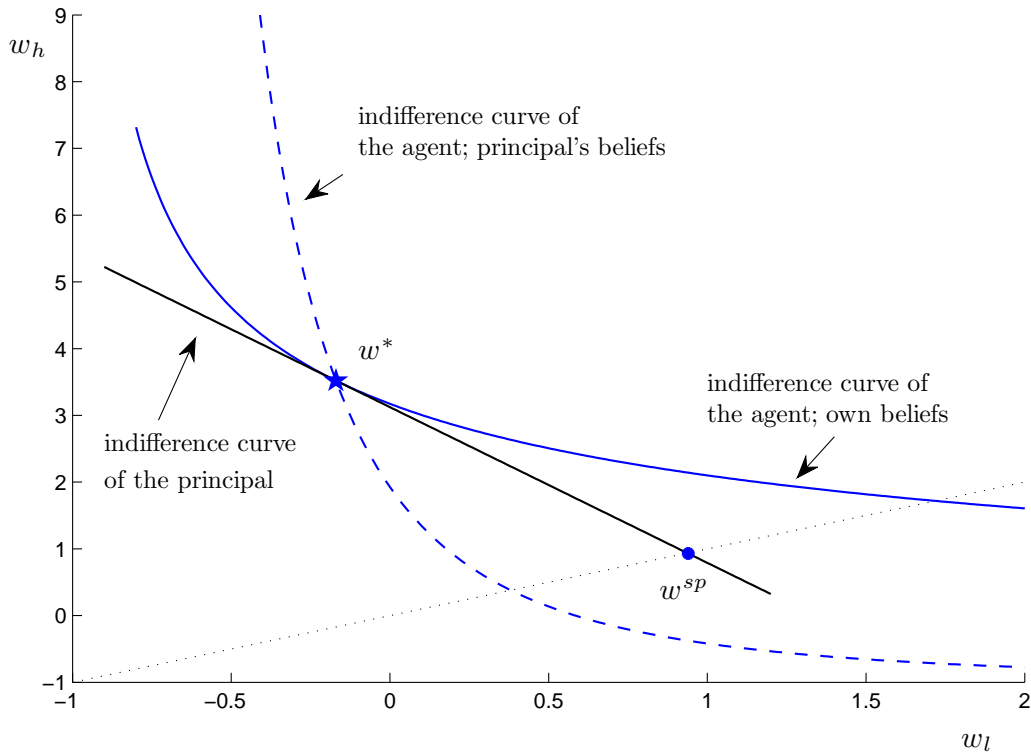


Figure 2: Dead Weight Loss from Arbitrage Under the Objectivists' View. The graph is built using the following parameter assumptions: $\pi_o = 1$, $\log(W + u_o) = 1$, $q_o = 0.7$, $f_o = 0.3$.

We next formally investigate second-best Pareto optimality of the equilibrium in our model. Our definition of second-best Pareto optimality assumes that the social planner has the same information and the same accounting technology as the principal. Thus, Pareto improvements are required to have wages that can screen out the desired agent type, and we do not consider allocations that rely on a menu of wages. We have in mind the following game: the social planner sets up the wage, agents learn their types and choose whether to apply for the managerial position; the manager is then randomly picked from those agents who apply. We require a second-best Pareto improvement to be implementable as a pure strategy Nash equilibrium in this game. While our definition of second-best Pareto optimality evaluates expected utilities prior to agents learning their types, if there is no Pareto improvement on this basis, there is no Pareto improvement immediately after agents learn their types.¹⁰

Definition 1. (*Second-best Pareto Optimality*) An allocation characterized by wage $w = (w_l, w_h)$ and hiring a type o (c) with probability p_o ($p_c = 1 - p_o$) is Second-best Pareto Optimal if there does not exist an alternative allocation with wage w' and hiring type o (c) with probability p'_o ($p'_c = 1 - p'_o$) that satisfies the following two conditions.

(i) All players weakly better off, with at least one player strictly better off:

$$\Pi(w', p'_o, p'_c) \geq \Pi(w, p_o, p_c) \text{ and } U_j(w', p'_o, p'_c) \geq U_j(w, p_o, p_c)$$

¹⁰There may be improvements after the hiring decision is committed to by all parties but the random payoff has not been realized, since this is a time at which we can improve risk-sharing without having to worry about self-selection constraints. For example, in equilibrium TS, the hired agent is known to be type c . Thus, once the agent is hired, Pareto improvement can be achieved by changing the wage to take more advantage of belief arbitrage (the new wage would be as specified in (7), but replacing f_m with f_c). This Pareto-dominant allocation, however, is not time-consistent. Note additionally that, in our model, agents not employed by the principal do not want to place bets on the project outcome because agents have the same beliefs as the principal about all other agents' abilities. The same results could be obtained, arguably more reasonably, if the agents believed they were at an informational disadvantage in evaluating other agents' abilities.

with at least one strict inequality, where the utility $U_j(w, p_o, p_c)$ of each individual agent j and the principal's profit $\Pi(w, p_o, p_c)$ are given by (9) and (10) respectively.

(ii) Hiring is incentive-compatible: $p'_o \in \{0, \pi_o, 1\}$, and

$p'_o = 0$ only if

$$q_c \log(W + w_h) + (1 - q_c) \log(W + w_l) \geq \log(W + u_c);$$

$$q_o \log(W + w_h) + (1 - q_o) \log(W + w_l) \leq \log(W + u_o);$$

$p'_o = 1$ only if

$$q_c \log(W + w_h) + (1 - q_c) \log(W + w_l) \leq \log(W + u_c);$$

$$q_o \log(W + w_h) + (1 - q_o) \log(W + w_l) \geq \log(W + u_o);$$

$p'_o = \pi_o$ only if

$$q_c \log(W + w_h) + (1 - q_c) \log(W + w_l) \geq \log(W + u_c);$$

$$q_o \log(W + w_h) + (1 - q_o) \log(W + w_l) \geq \log(W + u_o).$$

Except in degenerate cases, the social planner is not able to implement a Pareto improvement over the model's equilibrium. This is true even in the case of the objectivist social planner, despite the intuition developed in the example with only one agent. The problem is that agents' self-selection constraints keep the social planner from implementing strategies the planner views as better for them because they do not look better under their own beliefs. These results are formalized in the following theorem:

Theorem 2. *Whether we evaluate expected utilities of agents using their own beliefs (subjectivists' view, setting $q_c^w = q_c$ and $q_o^w = q_o$) or using the principal's beliefs (objectivists' view, setting $q_c^w = f_c$ and $q_o^w = f_o$), all equilibria are second-best Pareto optimal in the sense of Definition 1 with the following exceptions occurring when the principal is indifferent between two equilibria.*

1. Under the subjectivists' view, if the principal is indifferent between Equilibrium *NS* and either Equilibrium *TS* or Equilibrium *VP*, then Equilibrium *NS* Pareto-dominates Equilibria *TS* and *VP*.
2. Under the objectivists' view, if the principal is indifferent between Equilibrium *TS* and either Equilibrium *NS* or Equilibrium *VP*, then Equilibrium *TS* Pareto-dominates Equilibria *NS* and *VP*; and, if the principal is indifferent between Equilibrium *NS* and Equilibrium *VP*, then Equilibrium *NS* Pareto-dominates Equilibrium *VP*.

Proof. If the equilibrium is not second-best Pareto-optimal, then there exists a Pareto-dominant allocation with wage w' and probability $p'_o \in \{0, 1, \pi_o\}$ of hiring type o that maximizes the principal's expected profit $(p_o f_o + (1 - p_o) f_c)(a_h - w_h) + (1 - p_o f_o - (1 - p_o) f_c)(a_l - w_l)$ subject to the inequalities listed in item (ii) of Definition 1, and subject to $U_j(w, p_o, 1 - p_o) \geq U_j(w^*, p_o^*, 1 - p_o^*)$, where w^* is the equilibrium wage and p_o^* is the equilibrium probability of hiring type o (the last constraint follows from item (i) in Definition 1). In equilibrium, the principal solves the same problem, only without the last constraint (see the proof of Theorem 1). Therefore, if the solution to the principal's problem is unique, there is no second-best Pareto improvement. If, however, there are several solutions to the principal's problem, these solutions can always be ranked based on the expected utility they offer to the agent. Under the subjectivists' view, agents receive their reservation utilities in all equilibria but *NS*. Thus, equilibrium *NS* Pareto-dominates the rest of the equilibria. Under the objectivists' view, the hired agent receives expected utility below the reservation level (as evaluated by the social planner). The reduction in the expected utility is the smallest in equilibrium *TS* and the largest in equilibrium *VP*, as verified in the Appendix. \square

According to Theorem 2, the agents are better off in equilibrium NS than in equilibrium TS under the subjectivists' view, but they are better off in equilibrium TS than in equilibrium NS under the objectivists' view. Intuitively, the equilibrium wage is sensitive to performance, and so the social planner finds the same wage more valuable to type c than type o if the social planner uses the principal's beliefs. If however, the social planner uses agents' own beliefs, then the social planner will find the same wage more valuable to type o than type c . In equilibrium NS , type c receives the equilibrium wage with probability $\pi_c < 1$, while in equilibrium TS , type c receives the equilibrium wage with probability 1. Although the wages are different in the two equilibria, the differences in the beliefs used by the two views are large enough to reverse the order of preference as described in the theorem.

It may seem that the main result in Theorem 2 is rather extreme: despite the undesirability of belief arbitrage in the objectivists' view, second-best Pareto improvement is generally not feasible. This is because belief arbitrage in our model helps screen out agent types. A social planner who does not have better information than the principal about agents' types, cannot reduce arbitrage without hurting production and thus reducing the principal's profit. It is worth noting, however, that there are parameter values ruled out in our model setup for which it is second-best Pareto-optimal for the objectivist social planner to shut the market down (which is also ruled out by our definition of Pareto improvement in Definition 1). This happens, for example, if the principal finds hiring a type c agent not profitable, and hires a type o agent in equilibrium only because belief arbitrage is sufficiently attractive. Implementing such a Pareto improvement would still require the social planner to have a very good knowledge of parameters, allowing the social planner to conclude that shutting the market down would improve welfare.

IV Extensions

In our model of section II, the principal has only one position, there are two types of candidates, and there are many candidates of each type. In this section, we show that the results are robust to introducing additional types, multiple positions, and letting the competent type c be pessimistic about own ability.

IV.A Adding Pessimistic and Incompetent Agents

Introducing additional type p agents who are pessimistic and incompetent, and have reasonably high reservation wage ($u_p \geq u_c$ is sufficient but not necessary), would not affect the equilibria described in Theorem 1. So long as the participation constraint for the new type of agent is above the two equilibria depicted in Figure 1, the new type of agent will not apply for any wage contract that appears in Equilibria VP and TS. Because the principal finds the new agent more expensive and less able than the other agents, the principal will not alter the wage contracts to attract the new agent. Adding other types of agents, however, may have some effect on the equilibrium. For example, adding a moderately confident agent with sufficiently low reservation utility will result in this agent accepting all wage contracts that can possibly attract anyone else (if the participation constraint line of the new agent is well below the equilibria depicted in Figure 1). If the principal views the new agent's ability as sufficiently low, the principal does not want to hire the agent, and mechanisms other than wage structure would be required to screen out these agents.

IV.B Pessimistic Competent Type

Our model assumes that both agent types are more optimistic about their own ability than the principal. If we alter the model by assuming that $q_c < f_c < q_o$ (instead of $f_c < q_c < q_o$), the results change as follows. Because the principal in this version is more optimistic than the competent type about the competent type's ability, the optimal wage that attracts type c is flat: $w_l = w_h = u_c$. This implies that, first, there is no pooling equilibrium, and second, the presence of the overconfident type does not impact the wage of the competent type when the competent type is hired in equilibrium. The description of the equilibrium where the overconfident type is hired remains unchanged.

There are several reasons, however, to view the original case as more interesting. First, the results in the original case are more robust to introduction of effort, where flat compensation becomes suboptimal. Second, the principal may have a preference for hiring optimistic agents. For example, Bolton, Brunnermeier, and Veldkamp (2009) argue that optimistic agents make better managers because they are better at coordinating actions and motivating their subordinates.

IV.C Scarce Agents, Many Positions

When the principal has many positions and there is a shortage of agents, the wages that arise in equilibrium are similar to those highlighted in Figure 1, as we show next.

Suppose that at each time $t > 0$, the principal is approached by one agent. The principal believes that the agent is type o with probability π_o and type c with probability $1 - \pi_o$.

(thus, knowing the types of past applicants does not inform the principal about the type of the current applicant). The principal offers the agent a menu of wage schedules, and the agent then decides whether to apply for one of them. Given two agent types, we can assume without loss of generality that the menu of wages consists of at most two wage schedules: one to attract type o and one to attract type c .

If the principal wishes to hire every agent, the principal's objective is to find the least expensive pair of wage schedules that would attract both types. The principal's problem is similar to the principal's problem in section II, and thus leads to essentially the same wage schedules (we omit the formal derivation because it closely resembles that for Theorem 1). Specifically, depending on parameter values, the principal finds it optimal to either offer one wage schedule represented by point "NS" in Figure 1 to every agent, or offer a pair of wage schedules represented by points "VP" and "TS" in Figure 1, and let each agent choose which wage schedule to apply for. Hiring only one agent type is not an equilibrium in this setting because the principal could benefit by switching to hiring both types using one of the above wage menus (Assumption 1 guarantees feasibility of both wage menus). Specifically, hiring only type o (with wage schedule "VP") is dominated by hiring both types with wage schedules "VP" and "TS," and hiring only type c (with wage schedule "TS") is dominated by hiring both types with wage schedule "NS". Thus, there are two types of pure strategy symmetric equilibria:

NS At each time t , the principal hires the arrived agent with wage

$$(11) \quad \begin{aligned} w_h^* &= \exp(u_c) \left(\frac{f_m(1-q_c)}{(1-f_m)q_c} \right)^{q_c-1} - W, \\ w_l^* &= \exp(u_c) \left(\frac{f_m(1-q_c)}{(1-f_m)q_c} \right)^{q_c} - W. \end{aligned}$$

TS_VP At each time t , the principal offers a menu of two wages, w_o^* and w_c^* , given by

$$(12) \quad \begin{aligned} w_{oh}^* &= (W + u_o) \left(\frac{f_o(1 - q_o)}{(1 - f_o)q_o} \right)^{q_o - 1} - W, \\ w_{ol}^* &= (W + u_o) \left(\frac{f_o(1 - q_o)}{(1 - f_o)q_o} \right)^{q_o} - W; \end{aligned}$$

and

$$(13) \quad \begin{aligned} w_{ch}^* &= \exp \left\{ \frac{(1 - q_c) \log(W + u_o) - (1 - q_o) \log(W + u_c)}{q_o - q_c} \right\} - W, \\ w_{cl}^* &= \exp \left\{ \frac{q_o \log(W + u_c) - q_c \log(W + u_o)}{q_o - q_c} \right\} - W; \end{aligned}$$

If the agent that arrives at time t is type o , the agent applies for wage w_o^* and is hired; if the agent is type c , the agent applies for wage w_c^* and is hired.

As in the main model, which solution is realized depends on the model parameters. Because in both solutions the principal hires all agents, the principal chooses the solution with the lowest expected wage.

V Conclusion

This paper looks at compensation mechanisms that attract talented managers when the principal disagrees with the agents' assessments of their own ability. In contrast with the common belief that stronger performance sensitivity attracts more talented managers, we show that stronger performance sensitivity may instead attract overly-confident managers who are incompetent. The principal may prefer to hire the confident but incompetent manager if belief arbitrage is profitable enough for the principal to overcome the manager's lower expected ability. This may increase or decrease welfare depending on whether we take a subjectivist view that respects individual differences in beliefs or an objectivist

view that says that welfare should be based on utilities measured using the same objective probabilities for all agents. Except in degenerate cases, the social planner cannot improve on the equilibrium in these models, since the agents' own self-selection constraints defeat the social planner's attempts to intervene on their behalf.

Appendix

Proof of Theorem 1.

Proof. Consider a pooling solution that attracts both types of agents. This wage schedule must solve the following problem:

Problem 1.

$$\min_w (f_m w_h + (1 - f_m) w_l) \quad \text{subject to}$$

$$(14) \quad \log(W + u_o) \leq q_o \log(W + w_h) + (1 - q_o) \log(W + w_l),$$

$$(15) \quad \log(W + u_c) \leq q_c \log(W + w_h) + (1 - q_c) \log(W + w_l),$$

$$(16) \quad 0 \leq w_h - w_l,$$

where $f_m \equiv f_o \pi_o + f_c \pi_c$. In this problem, the constraints (14) and (15) insure that the applying for the position is incentive-compatible for both types. The wage schedule $w^* = (w_l^*, w_h^*)$ that minimizes the objective subject only to constraint (15) can be found by solving the following system of equations (the first order conditions to the problem with respect to w_l and w_h , combined with condition (15) satisfied with equality):

$$\begin{aligned}
f_m &= \lambda q_c \frac{1}{W + w_h}; \\
1 - f_m &= \lambda(1 - q_c) \frac{1}{W + w_l}; \\
\log(W + u_c) &= q_c \log(W + w_h) + (1 - q_c) \log(W + w_l);
\end{aligned}$$

Wage schedule w^* given by (7) solves the above system of equations. Moreover, it also satisfies (14) and (16) with strict inequality. Since (14) and (16) are not binding at any solution, this is a maximization of a nonconstant linear objective function with a strictly convex constraint set, so the solution is unique.

Consider next a separating solution that attracts only type o but not type c . This wage must solve the following problem:¹¹

Problem 2.

$$\min_w (f_o w_h + (1 - f_o) w_l) \quad \text{subject to}$$

$$(17) \quad \log(W + u_o) \leq q_o \log(W + w_h) + (1 - q_o) \log(W + w_l),$$

$$(18) \quad \log(W + u_c) \geq q_c \log(W + w_h) + (1 - q_c) \log(W + w_l),$$

$$(19) \quad 0 \leq w_h - w_l,$$

where the constraints (17) and (18) insure that the applying for the position is incentive-compatible only for type o . The wage schedule $w^* = (w_l^*, w_h^*)$ that minimizes the objective subject only to constraint (17) can be found by solving the following system of equations

¹¹Note that some agents are indifferent between applying and not applying if either (17) or (18) is satisfied with equality. This is not a problem, however, which can be seen following the logic of footnote 6.

(the first order conditions to the problem with respect to w_l and w_h , combined with condition (17) satisfied with equality):

$$\begin{aligned} f_o &= \lambda q_o \frac{1}{W + w_h}; \\ 1 - f_o &= \lambda(1 - q_o) \frac{1}{W + w_l}; \\ \log(W + u_o) &= q_o \log(w_h + W) + (1 - q_o) \log(W + w_l); \end{aligned}$$

Wage schedule w^* given by (3) solves the above system of equations. Moreover, it also satisfies (18) and (19). Since (18) and (19) are not binding at any solution, this is a maximization of a nonconstant linear objective function with a strictly convex constraint set, so the solution is unique.

Consider finally a separating solution that attracts only type o but not type c . This wage must solve the following problem:

Problem 3.

$$\min_w (f_c w_h + (1 - f_c) w_l) \quad \text{subject to}$$

$$(20) \quad \log(W + u_o) \geq q_o \log(W + w_h) + (1 - q_o) \log(W + w_l),$$

$$(21) \quad \log(W + u_c) \leq q_c \log(W + w_h) + (1 - q_c) \log(W + w_l),$$

$$(22) \quad 0 \leq w_h - w_l,$$

where the constraints (20) and (21) insure that the applying for the position is incentive-compatible only for type c . The solution $w^* = (w_l^*, w_h^*)$ to Problem 1 satisfies the following first order conditions with respect to w_h and w_l and complementary slackness conditions:

$$\begin{aligned}
f_c &= \lambda_1 q_c \frac{1}{W + w_h} - \lambda_2 q_o \frac{1}{W + w_h}; \\
1 - f_c &= \lambda_1 (1 - q_c) \frac{1}{W + w_l} - \lambda_2 (1 - q_o) \frac{1}{W + w_l}; \\
\log(W + u_o) &= q_o \log(W + w_h) + (1 - q_o) \log(W + w_l); \\
\log(W + u_c) &\leq q_c \log(W + w_h) + (1 - q_c) \log(W + w_l); \\
0 &= \lambda_3 (w_h - w_l); \\
0 &\leq \lambda_k, \quad k=1,2,3.
\end{aligned}$$

Wage schedule w^* given by (3), combined with

$$\begin{aligned}
\lambda_1 &= \exp\left(\frac{q_o \log(W + u_c) - q_c \log(W + u_o)}{q_o - q_c}\right) \\
&\quad \times \frac{[q_o(1 - f_c) - (1 - q_o)f_c \exp((\log(W + u_o) - \log(W + u_c))/(q_o - q_c))]}{q_o - q_c}, \\
\lambda_2 &= \exp\left(\frac{q_o \log(W + u_c) - q_c \log(W + u_o)}{q_o - q_c}\right) \\
&\quad \times \frac{[q_c(1 - f_c) - (1 - q_c)f_c \exp((\log(W + u_o) - \log(W + u_c))/(q_o - q_c))]}{q_o - q_c},
\end{aligned}$$

and $\lambda_3 = 0$, satisfies the above system of equations ($\lambda_1 > 0$ and $\lambda_2 > 0$ follows from the assumption 1c). Since Problem 3 is a maximization of a nonconstant linear objective function with a strictly convex constraint set, w^* is the unique solution to this problem.

Feasibility of all three solutions for some parameter values is illustrated in Table 1. \square

Table 1: Numerical Illustration of Possible Equilibria.

The numerical examples reported in this table illustrate that, depending on parameter values, any set of strategies described in Theorem 1 (VP, TS, and NS) can be a solution. All of the examples assume $q_o = 0.7$, $q_c = 0.6$, $\log(W + u_o) = 1$, $\log(W + u_c) = 0.9$, $f_o = 0.1$, $f_c = 0.2$, $W = 1$, $a_l = 5$. Each line specifies the remaining parameters π_o and a_h and reports the principal's payoffs Π_o from hiring only type o , Π_c from hiring only type c , and Π_{ns} from hiring both types (note that the assumed parameters satisfy assumptions 1a - 1c). The solution corresponds to the highest value of the principal's payoff.

	Parameters		Principal's Payoff			Solution
	π_o	a_h	Π_{vp}	Π_{ns}	Π_{ts}	
Example I	0.5	10	5.53	5.30	5.19	VP
Example II	0.5	15	6.03	6.05	6.19	TS
Example II	0.1	15	6.03	6.26	6.19	NS

Lemma 1. *The following inequalities hold,*

$$\frac{(1 - f_o)q_o}{(1 - q_o)f_o} > \frac{(1 - f_m)q_c}{(1 - q_c)f_m} > \frac{(1 - f_c)q_c}{(1 - q_c)f_c} > 1.$$

Proof. The first inequality holds because, according to Assumption 1(a), $q_o > q_c$, and $f_o < f_m = \pi_o f_o + \pi_c f_c$, which implies that $(1 - f_o) > (1 - f_m)$. Similarly, the second inequality holds because, according to Assumption 1(a), $f_m < f_c$. Finally, the last inequality holds because, from Assumption 1(a), $(1 - f_c) > (1 - q_c)$ and $q_c > f_c$. \square

Proof of Theorem 2.

Proof. To complete the proof offered in the text, it remains to show that $U_{TS} > U_{NS} > U_{VP}$, where U_{EQ} is the agents' expected utility in equilibrium $EQ \in NS, VP, TS$, evaluated using (9) with the corresponding equilibrium hiring probabilities, equilibrium wage, and objectivists' beliefs about ability given by $q_c^w = f_c$ and $q_o^w = f_o$. To focus on the parts of the expected utility expressions that differ across equilibria, let

$$(23) \quad u_{EQ} = N [U_{EQ} - \pi_c \log(W + u_c) - \pi_o \log(W + u_o)].$$

Substituting (9), $q_c^w = f_c$, and $q_o^w = f_o$ into (23), we obtain

$$(24) \quad \begin{aligned} u_{EQ} = & -p_c \log(W + u_c) - p_o \log(W + u_o) \\ & + p_c [f_c \log(W + w_h^*) + (1 - f_c) \log(W + w_l^*)] \\ & + p_o [f_o \log(W + w_h^*) + (1 - f_o) \log(W + w_l^*)]. \end{aligned}$$

When $EQ = VP$, we have $p_o = 1$, $p_c = 0$, and w^* is given by (3). Thus, (24) becomes

$$\begin{aligned}
u_{VP} &= -\log(W + u_o) + f_o \log(W + w_h^*) + (1 - f_o) \log(W + w_i^*) \\
&= -\log(W + u_o) + f_o \log \left((W + u_o) \left(\frac{f_o(1 - q_o)}{(1 - f_o)q_o} \right)^{q_o - 1} \right) \\
&\quad + (1 - f_o) \log \left((W + u_o) \left(\frac{f_o(1 - q_o)}{(1 - f_o)q_o} \right)^{q_o} \right) \\
&= [q_o - f_o] \log \left(\frac{f_o(1 - q_o)}{(1 - f_o)q_o} \right).
\end{aligned}$$

Similarly, when $EQ = NS$, we have $p_o = \pi_o$, $p_c = \pi_c$, and w^* is given by (7). Thus, recalling from the text before (7) that $f_m \equiv f_o \pi_o + f_c \pi_c$, (24) becomes

$$\begin{aligned}
u_{NS} &= -\pi_c \log(W + u_c) - \pi_o \log(W + u_o) + f_m \log(W + w_h^*) + (1 - f_m) \log(W + w_i^*) \\
&= (\pi_o - 1) \log(W + u_c) - \pi_o \log(W + u_o) + f_m \log \left((W + u_c) \left(\frac{f_m(1 - q_c)}{(1 - f_m)q_c} \right)^{q_c - 1} \right) \\
&\quad + (1 - f_m) \log \left((W + u_c) \left(\frac{f_m(1 - q_c)}{(1 - f_m)q_c} \right)^{q_c} \right) \\
&= \pi_o [\log(W + u_c) - \log(W + u_o)] + [q_c - f_m] \log \left(\frac{f_m(1 - q_c)}{(1 - f_m)q_c} \right).
\end{aligned}$$

Finally, when $EQ = TS$, we have $p_o = 0$, $p_c = 1$, and w^* is given by (5). Thus, (24) becomes

$$\begin{aligned}
u_{TS} &= -\log(W + u_c) + f_c \log(W + w_h^*) + (1 - f_c) \log(W + w_i^*) \\
&= -\log(W + u_c) + \frac{f_c(1 - q_c) \log(W + u_o) - f_c(1 - q_o) \log(W + u_c)}{q_o - q_c} \\
&\quad + \frac{(1 - f_c)q_o \log(W + u_c) - (1 - f_c)q_c \log(W + u_o)}{q_o - q_c} \\
&= -\log(W + u_c) + \frac{[f_c - q_c] \log(W + u_o) - [f_c - q_o] \log(W + u_c)}{q_o - q_c} \\
&= \frac{-[q_o - q_c] \log(W + u_c) + [f_c - q_c] \log(W + u_o) - [f_c - q_o] \log(W + u_c)}{q_o - q_c} \\
&= [f_c - q_c] \frac{[\log(W + u_o) - \log(W + u_c)]}{q_o - q_c}.
\end{aligned}$$

The above expressions can further be rewritten as

$$(25) \quad u_{VP} = -[q_o - f_o] \log \left(\frac{(1 - f_o)q_o}{f_o(1 - q_o)} \right),$$

$$(26) \quad u_{NS} = -\pi_o[\log(W + u_o) - \log(W + u_c)] - [q_c - f_m] \log \left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)} \right),$$

$$(27) \quad u_{TS} = -[q_c - f_c] \frac{\log(W + u_o) - \log(W + u_c)}{q_o - q_c}.$$

We show that $u_{TS} > u_{NS}$ as follows.

$$\begin{aligned} u_{TS} &= -[q_c - f_c] \frac{\log(W + u_o) - \log(W + u_c)}{q_o - q_c} \\ &> -[q_c - f_c] \log \left(\frac{(1 - f_c)q_c}{f_c(1 - q_c)} \right) \\ &> -[q_c - f_c] \log \left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)} \right) \\ &> -[q_c - f_m] \log \left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)} \right) \\ &> -\pi_o[\log(W + u_o) - \log(W + u_c)] - [q_c - f_m] \log \left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)} \right) \\ &= u_{NS}. \end{aligned}$$

In the above, the first inequality follows from Assumption 1(a) which implies that $q_c - f_c > 0$ and from Assumption 1(c) which states that:

$$\frac{\log(W + u_o) - \log(W + u_c)}{q_o - q_c} < \log \left(\frac{(1 - f_c)q_c}{f_c(1 - q_c)} \right).$$

The second inequality follows from Lemma 1, and $q_c - f_c > 0$ (as already noted). The third inequality follows from Assumption 1(a) implying that $q_c > f_c > f_m$ and Lemma 1 implying that $\log \left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)} \right) > \log(1) > 0$. The last inequality follows from Assumption 1(d) implying that $\log(W + u_o) - \log(W + u_c) > 0$.

We show that $u_{NS} > u_{VP}$ as follows.

$$\begin{aligned}
u_{NS} &= -\pi_o[\log(W + u_o) - \log(W + u_c)] - [q_c - f_m] \log\left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)}\right) \\
&> -\pi_o[q_o - q_c] \log\left(\frac{(1 - f_c)q_c}{f_c(1 - q_c)}\right) - [q_c - f_m] \log\left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)}\right) \\
&> -[q_o - q_c] \log\left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)}\right) - [q_c - f_m] \log\left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)}\right) \\
&= -[q_o - f_m] \log\left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)}\right) \\
&> -[q_o - f_o] \log\left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)}\right) \\
&> [q_o - f_o] \log\left(\frac{(1 - f_o)q_o}{f_o(1 - q_o)}\right) \\
&= u_{VP}.
\end{aligned}$$

In the above, the first inequality follows from Assumption 1(c) and Assumption 1(a) (which implies $q_o > q_c$). The second inequality obtains because, from Assumption 1(a), $\pi_o < 1$, and again $q_o > q_c$. The third inequality follows from Assumption 1(a) implying that $q_o > f_m > f_o$, and Lemma 1 implies that $\log\left(\frac{(1 - f_m)q_c}{f_m(1 - q_c)}\right) > \log\left(\frac{(1 - f_c)q_c}{f_c(1 - q_c)}\right) > \log(1) = 0$. Finally, the last inequality follows from Lemma 1 and Assumption 1(a) which implies $q_o > f_o$. Thus, we have $u_{TS} > u_{NS} > u_{VP}$. \square

References

- Adrian, Tobias, and Mark M. Westerfield. 2008. "Disagreement and Learning in a Dynamic Contracting Model." Federal Reserve Bank of New York Staff Report, 269.
- Arak, Marcelle, Arturo Estrella, Laurie Goodman, and Andrew Silver. 1988. "Interest Rate Swaps: An Alternative Explanation." *Financial Management*, 17(2): 12-18.
- Ben-David, John R. Graham, and Campbell R. Harvey. 2007. "Managerial Overconfi-

dence and Corporate Policies.” NBER working paper No. 16215.

Bhattacharya, S., and P. Pfleiderer. 1985. “Delegated Portfolio Management.” *Journal of Economic Theory*, 36(1): 1-25.

Bolton, Patrick, Markus K. Brunnermeier, and Laura Veldkamp, 2009. ”Leadership, Coordination and Mission-Driven Management.” Working paper, NYU.

Dybvig, Philip H., Heber K. Farnsworth, and Jennifer N. Carpenter. 2010. “Portfolio Performance and Agency.” *Review of Financial Studies*, 23(1): 1-23.

Dybvig, Philip H., and L. C. G. Rogers. 2010. “High Hopes and Disappointment.” Working paper, Washington University in St. Louis.

Gervais, Simon, J.B. Heaton, and Terrance Odean. 2010. “Overconfidence, Compensation Contracts, and Capital Budgeting.” Working paper, Duke University.

Kozzegi, Botond, and Wei Li. 2008. “Drive and Talent.” *Journal of the European Economic Association*, 6(1): 210-236.

Khorana, Ajay , Henri Servaes, and Lei Wedge. 2007. “Portfolio manager ownership and fund performance.” *Journal of Financial Economics*, 85: 179-204.

Landier, Augustin, and David Thesmar. 2009. “Financial Contracting with Optimistic Entrepreneurs.” *Review of Financial Studies*, 22(1): 117-150.

Malmendier, Ulrike, Stefan Nagel. 2011. “Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?” *Quarterly Journal of Economics*, February, 126(1): 373-416.

Malmendier, Ulrike, Geoffrey Tate. 2005. "Does Overconfidence Affect Corporate Investment? CEO Overconfidence Measures Revisited." *European Financial Management*, 11(5): 649-659.

Malmendier, Ulrike, Geoffrey Tate, Jun Yan. 2011. "Overconfidence and Early-life Experiences: The Effect of Managerial Traits on Corporate Financial Policies." *Journal of Finance*, 66(5): 1687-1733.

Manove, Michael, and A. Jorge Padilla. 1999. "Banking (conservatively) with optimists." *RAND Journal of Economics*, 30(2): 324-350.

Riley, John G. 2001. "Silver Signals: Twenty-Five Years of Screening and Signaling." *Journal of Economic Literature*, 39(2): 432-478.

Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics*, 87(3): 355-374.

Van den Steen, E. 2004. "Rational Overoptimism (and Other Biases)" *American Economic Review*, 94(4): 1141-1151.

Van den Steen, E. 2005. "Organizational Beliefs and Managerial Vision." *Journal of Law, Economics, and Organization*, 21(1): 256-283.