Essays on Optimal Contracts with Overconfidence

Justin Downs

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Fahad Khalil, Chair

Jacques Lawarree, Chair

Xu Tan

Program Authorized to Offer Degree:
Economics

University of Washington

Abstract

Essays on Optimal Contracts with Overconfidence

Justin Downs

Chair of Supervisory Committee:
Professor Fahad Khalil
Department of Economics

Professor Jacques Lawarree
Department of Economics

This dissertation studies the effect of overconfidence on markets and organizations with asymmetric information. In the first chapter, I introduce overconfidence into a standard information gathering contracting model. A principal (she) hires an agent (he) to gather information about a project's cost before he implements the project, and the agent overestimates the probability of having a low implementation cost. The agent's overconfidence makes him more willing to sign the contract, but less willing to gather information, and increases in overconfidence may increase or decrease the principal's profit. In the second chapter, I study a labor market where firms hire overconfident workers who have private information about their productivity. I derive the optimal contracts for both a monopsonistic market, where one firm makes take-it-or-leave-it offers to the workers, as well as a competitive market, where many firms compete for the services of workers. Overconfidence causes the optimal contract to be distorted away from the efficient outcome in both markets, but a monopsonistic firm internalizes these distortions while a competitive firm does not. The main result is that monopsonistic markets can be more efficient than competitive markets. In the third chapter, I provide a review of several mathematical definitions of overconfidence used in the contract theory literature and apply them all to a generalized version of the information gathering model from Chapter 1. The effects overconfidence has on the agent's willingness to participate, to gather information, and on the principal's profit are all sensitive to the mathematical definition of overconfidence used in the model.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Alexander Rodivilov, Richard Van Weelden, Quan Wen, and especially Chris Anderson and Xu Tan for their valuable comments and insights which helped me develop and refine this dissertation, as well as helpful discussion from participants in seminars at UW and the Midwest Economic Theory Conference 2019. I would also like to thank Simon Reeve-Parker and the other staff members in the UW Econ department who provided me with extensive administrative support. I am grateful for the many friends and family members who helped me in ways small and large, direct and indirect, noticeable and subtle, to live up to my potential and achieve a level of personal fulfillment that I have never had before. I am also grateful for generous financial support from the Grover and Creta Ensley Fellowship.

Finally, I am extremely thankful to Fahad Khalil and Jaques Lawarree. As co-chairs of my dissertation committee and primary research advisors, they taught me to distill my often chaotic and unfocused thought-process into meaningful economic concepts and intuition. They also supported, encouraged, and motivated me at many times when I needed it most. Their mentorship has been invaluable to me.

Chapter 1

Information Gathering by Overconfident Agents

Abstract: A principal hires an agent to both gather information about a project's costs and implement it. The agent's information gathering effort and what he learns are his private information. I allow the agent to be overconfident in the sense that he underestimates his expected cost of implementation and study the effects this overconfidence has on the efficiency of information acquisition and implementation. Overconfidence makes the agent more willing to accept a given contract but may dampen his incentive to gather information. As a result, information may not be gathered in equilibrium due solely to the agent's overconfidence which causes inefficiencies in the project's implementation as well. When the agent's information gathering cost is low enough, the principal's payoff is non-monotonic in the degree of overconfidence, increasing for both low and high levels of overconfidence, but decreasing for intermediate levels.

1.1 Introduction

The efficient acquisition and distribution of information is critical to the success of any organization. In many cases, when embarking on a new project, the decision of whether to gather planning information must be delegated to some member of the organization by its leader. For example, a firm owner needs its manager to obtain information on the profitability of new business projects or product lines, politicians employ advisors to assess the quality of campaign strategies, and professional sports teams send scouts to evaluate the skills and talent of potential recruits. When pre-project planning and project implementation must be done by the same individual or team, incentives for efficient planning and implementation are intertwined, and the efficiency of one, the other, or both may be affected. Further complications may arise when the members of the organization disagree about what is likely to be learned from pre-project planning. This raises the question: how might disagreement affect the structure of contracts that govern both planning and implementation? Due to its prevalence in the real world and in the behavioral contract theory literature, in this paper I will focus on the case of an overconfident agent.[1]

Suppose the leader of an organization (principal) needs to provide incentives to a subordinate (agent) for both acquiring and revealing information about costs before implementing a project. The agent is overconfident about *the distribution of his implementation cost*: relative to the principal, he overestimates the prior probability that his implement cost will be low.[2] Overconfidence has two effects on the agent's decision-making. First, he overestimates his expected payoff from any contract that induces information gathering. This makes him relatively

[1] For empirical evidence of overconfidence, see Camerer and Lovallo (1999), Barber and Odean (2000), Biais et al (2005), Malmendier and Tate (2005,2008), DellaVigna and Malmendier (2006), DellaVigna (2009), Larkin and Leider (2012), Ben-David et al (2013), Sautmann (2013), and Pikulina et al (2017).

[2] This definition of overconfidence has been used before in moral hazard settings. See, for example, Santos-Pinto (2008) or De La Rosa (2011).

more willing to accept the contract, allowing the principal to induce his participation with a lower expected wage. I refer to this as the *participation effect* of overconfidence. Second, the agent overestimates his expected payoff from remaining uninformed and implementing the project without knowledge of his cost. This makes shirking on information gathering relatively more attractive and so more powerful incentives are required to induce an overconfident agent to gather information. The increased power of incentives needed to induce information gathering is what I refer to as the *incentive effect* of overconfidence.

The participation and incentive effects interact to influence the efficiency of both information acquisition and implementation, as well the principal's payoff from the optimal contract. Suppose that, in the benchmark case where the principal and agent agree (share common priors), the optimal contract achieves the first best outcome for the principal. In this contract the agent gathers information, implementation is efficient, and the principal's profit equals her maximum expected surplus from trade. Under this assumption, any changes in the principal's payoff or inefficiencies in implementation or information acquisition can be directly attributed to disagreement between the principal and agent.

When the agent's degree of overconfidence is low, the participation effect is dominant as the effect overconfidence has on information gathering incentives is minimal. Information acquisition and project implementation are both efficient, and the principal's payoff is increasing in the agent's overconfidence. As overconfidence increases, the incentive effect grows relative to the participation effect and some novel results emerge. First, there comes a point at which the increased cost of information gathering incentives outweighs the benefit of that information itself. In this case, the optimal contract will not induce the agent to gather information and implementation will necessarily be inefficient. Second, even when it remains optimal to induce information gathering, the principal's payoff from the optimal contract may start to decrease in the degree of overconfidence. Thus, the principal's payoff may exhibit non-monotonicity in the agent's overconfidence. Finally, when the agent's overconfidence is extreme, rewards meant to prevent the agent from overstating his costs conflict with rewards meant to encourage the agent to gather information to such an extent that information gathering is infeasible altogether. Nevertheless, the principal's payoff will again be increasing in the degree of overconfidence, as the payment required to induce the agent's participation is decreasing in his expected cost, which decreases with overconfidence.

These results illustrate the main ideas of this paper. First, even with a principal providing explicit incentives to do so, it may be difficult or impossible to motivate an overconfident agent to learn. This could result in, for example, excess entry into new markets (Camerer and Lovallo (1999)), overzealousness in corporate acquisitions (Malmendier and Tate (2008)), or distortions in corporate investments (Malmendier and Tate (2005)). And second, the additional costs associated with inducing information gathering may outweigh other benefits of hiring an overconfident agent. This calls into question the wisdom of tournament-style promotion processes that may select for overconfident managers or CEOs (Goel and Thakor (2008), Park and Santos-Pinto (2010)) as well as the hiring of experts who may enjoy lower implementation or information acquisition costs as well as increased levels of overconfidence (See Gervais, Heaton, and Odean (2011) for citations on overconfident experts).

In addition to the baseline model with an overconfident agent, I also consider the case of diffidence, where the agent *overestimates* his expected implementation cost relative to the principal. As with overconfidence, there is both a participation and incentive effect of diffidence. Unlike the case of overconfidence, both of these effects work against the principal. Diffidence makes the agent *undervalue* contracts that induce information gathering, making it more expensive for the principal to induce his participation. Diffidence also makes more attractive the deviation where the agent remains uninformed and claims to have learned that his costs are high, producing a *negative* incentive effect that is analogous to that of the case of overconfidence. I find that the principal's profit is unambiguously decreasing as the agent becomes more diffident, which is consistent with previous results in the literature.

This paper contributes to the literatures on information gathering and contracting between parties with non-common priors. As for the information gathering literature, researchers have uncovered a variety of factors that influence whether it is either efficient or optimal for a principal to incent or deter information gathering by an agent.[3] One common theme of this literature is that high (resp. low) powered incentives will be used to incent (resp. deter) information gathering, as in, for example, Crémer et al. (1998a), Lewis and Sappington (1997), Dai et al. (2006), and Szalay (2009). This is a pattern I find as well, but for a new reason: for *any* positive information gathering cost, the power of incentives needed to induce information gathering grows as the degree of overconfidence grows.

The previous literature on contracting under non-common priors also puts its emphasis on overconfidence, but typically finds that the principal benefits from the agent's overconfidence. In many cases, these benefits resemble the participation effect introduced above.[4] In other cases, the benefits come from increased effort by agents or equilibrium contracts that exploit the agent's

[3] These factors include the timing of information gathering (Crémer and Khalil (1994), Crémer et al. (1998b), Kessler (1998), Terstiege (2012)), whether information gathering is productive or strategic (Crémer and Khalil (1992), Crémer et al. (1998b), Schmitz (2008), Hoppe and Schmitz (2010), Terstiege (2016)), the number of agents (Crémer and Khalil (1992), Compte and Jehiel (2008)), as well as whether to delegate, integrate or separate planning and implementation (Lewis and Sappington (1997), Khalil et al. (2006), Dai et al. (2006), Shin and Yun (2008), Hoppe and Schmitz (2013)).
[4] See Santos-Pinto (2008), De La Rosa (2011), and Sautmann (2013).

mispredictions about the future.[5] My work contributes to this literature by revealing a possible cost of overconfidence related to the provision of incentives and showing that a principal may not always benefit from increases in the agent's overconfidence.

There are three other papers which also consider models of asymmetric information that include both information gathering and overconfidence. They are Goel and Thakor (2008), Gervais, Heaton and Odean (2011), and Iossa and Martimort (2015). The first two papers study investment decisions made by risk averse and overconfident CEOs, where overconfidence is modeled as an overestimation of the correlation between a signal privately observed by the CEO and the true underlying profitability of an investment opportunity.[6] In Iossa and Martimort (2015), the agent may be overconfident ("optimistic", in their terminology) in that, conditional on his information gathering effort, he overestimates the probability of becoming informed. In each of these papers, overconfidence motivates the agent to exert relatively more information gathering effort than if he shared the principal's belief, which reduces the expected cost to the principal of inducing a given level of effort. As such, the amount of information gathered in equilibrium is increasing in the agent's overconfidence. In contrast, I show that overconfidence may reduce the amount of information that is obtained in equilibrium. This stems from the fact that in each of those papers the agent is overconfident about the *efficiency* of learning, which encourages him to gather information, while in my model the agent is overconfident in the *outcome* of learning, which can discourage him from gathering information.

The remainder of the paper is organized as follows. In Section 2, I present the baseline information gathering model and analysis for the case of an overconfident agent. Section 3

---

[5] Overconfidence increases equilibrium effort in Fang and Moscarini (2005) and Gervais and Goldstein (2007). Contracts exploit the agent's mispredictions in Eliaz and Spiegler (2006) and Grubb (2009).
[6] This form of overconfidence may also be thought of as over-precision.

considers the case of a diffident agent, while Section 4 concludes. Proofs for Sections 2 and 3 can be found in Appendices A and B, respectively.

1.2 Model

A principal (she) hires an agent (he) to implement a project of size q. For simplicity, let $q \in \{0,1\}$, so that the project is either implemented ($q = 1$) or shut down ($q = 0$).[7] The principal values the project according to the function $V(q) = Vq$. The agent's cost of production is βq where $\beta \in \{\beta_L, \beta_H\}$, $\beta_H > \beta_L > 0$ and $\Delta\beta \equiv \beta_H - \beta_L$[8]. Project returns are such that it is *ex post* efficient to implement the project if and only if costs are low: $\beta_H > V > \beta_L > 0$. Thus, efficient implementation requires information about costs, which makes that information valuable *ex ante*.

At the time of contracting, neither the principal nor the agent knows the value of β. After the contract is signed the agent may exert effort $a \in \{0,1\}$ at utility cost ca to learn the value of β with probability a, where $c \geq 0$ is the incremental cost of information gathering effort. Importantly, whether the agent exerts effort or not and what the agent learns, if anything, are each the agent's private information. The principal compensates the agent for his costs of implementation and information gathering effort with a monetary transfer t. The payoffs of the principal and agent are:

$$\Pi(q, t) = Vq - t$$

$$U(q, t; \beta) = t - \beta q - ca$$

respectively. Each player has a reservation payoff that is normalized to zero.

---

[7] If the project size is binary, i.e., $q \in \{\underline{q}, \overline{q}\}$ where $\overline{q} > \underline{q} \geq 0$, letting $\overline{q} = 1$ and $\underline{q} = 0$ is without loss of generality.

[8] In Section 2.b., I briefly discuss a more general model where the projects' size can be varied continuously, i.e., $q \in [0, \infty)$. All the main results carry through in this extension. See Downs (2019) for details.

Prior to contracting, the principal and agent are each endowed with beliefs about the distribution of the implementation cost β. The key assumption of this paper is that the principal and agent may hold different beliefs about the distribution of β. Specifically, let $\eta \in (0,1)$ (resp. $\tilde{\eta} \in (0,1)$) be the principal's (resp. agent's) prior belief that $\beta = \beta_L$. Then, the principal's subjective expected cost is $\beta_P \equiv \eta\beta_L + (1-\eta)\beta_H$ and the agent's subjective expected cost is $\beta_A \equiv \tilde{\eta}\beta_L + (1-\tilde{\eta})\beta_H$. In Section 2, I will assume that $\tilde{\eta} \geq \eta$, so that the beliefs characterize an agent who is *overconfident relative to the principal* (overconfident hereafter) about his cost of production.[9] I will also assume that these beliefs are common knowledge and that the principal and agent *agree to disagree*, i.e., neither principal nor agent will update their belief about β unless information gathering has taken place and its outcome revealed.[10] The timing of the game is as follows:

- $t = 1$: The principal offers a menu of contracts $(t_i, q_i)_{i\in\{L,H\}}$ to the agent, who accepts or rejects. If he rejects, the game ends and the players receive their reservation payoffs. If he accepts, the game proceeds to $t = 2$.
- $t = 2$: The agent privately exerts information gathering effort a. He privately learns his type β if and only if $a = 1$.
- $t = 3$: The agent implements a project of size q, which is verifiable and delivered directly to the principal.
- $t = 4$: Contractual payments are made, and payoffs are realized.

---

[9] See Section 3.a. for the case of diffidence ($\tilde{\eta} \leq \eta$).
[10] This assumption allows me to focus on the implications of differences in beliefs, rather than their source. It has been used extensively in models of non-common priors. See Eliaz and Spiegler (2006,2007,2008), Gervais and Goldstein (2007), Goel and Thakor (2008), Santos-Pinto (2008), Grubb (2009), De La Rosa (2011), Gervais, Heaton, and Odea (2011), Sautmann (2013), or Iossa and Martimort (2015).

1.2.a Costless Information Gathering

The case where information gathering is costless $(c = 0)$ is useful for highlighting one of the effects of overconfidence, which I refer to as the *participation effect*. Since becoming informed comes at no cost to the agent, and there is no strategic benefit to remaining uninformed, there is no loss of generality in assuming the agent gathers information at the principal's request. The situation simplifies to a model of adverse selection with *ex ante* contracting and non-common priors. The principal's maximization problem is $[P_0]$:

$$\Pi_0(\tilde{\eta}) = \max_{t_H, t_L, q_H, q_L} \eta[Vq_L - t_L] + (1-\eta)[Vq_H - t_H] \qquad \text{subject to}$$

$$\tilde{\eta}[t_L - \beta_L q_L] + (1-\tilde{\eta})[t_H - \beta q_H] \geq 0 \qquad (\boldsymbol{PC})$$

$$t_L - \beta_L q_L \geq t_H - \beta_H q_H \qquad (\boldsymbol{IC_L})$$

$$t_H - \beta_H q_H \geq t_L - \beta_H q_L \qquad (\boldsymbol{IC_H})$$

where (PC) is the *ex-ante* participation constraint and $(IC_L), (IC_H)$ are truth-telling constraints for each realization of β.[11] Under agreement $(\tilde{\eta} = \eta)$, the solution to this problem is a contract where implementation is efficient $(q_L = 1, q_H = 0)$ and the principal's payoff is equal to the maximum expected total surplus $\eta[V - \beta_L]$.

When beliefs differ $(\tilde{\eta} > \eta)$ the agent's overconfidence gives the principal an incentive to structure a "side-bet" into the contract that pays off for the agent when he implements the project, i.e., when his realized cost of production is low, the state which the agent deems more likely (relative to the principal). This side-bet takes the form of an increase in t_L and a corresponding decrease in t_H that leaves the agent indifferent but generates profits for the principal. For instance, hold (q_L, q_H) fixed and take a pair (t_L, t_H) that strictly satisfies all the constraints. The principal

[11] I refer to these as "truth-telling constraints" to distinguish them from "information gathering incentive constraints" that will be introduced in Section 2.b.

can reduce t_H by some $\delta > 0$ and increase t_L by $\frac{1-\tilde{\eta}}{\tilde{\eta}}\delta$, which leaves the agent indifferent while increasing the principal's payoff by $\frac{\tilde{\eta}-\eta}{\tilde{\eta}}\delta > 0$. Evidently, the principal has no such betting motive when $\tilde{\eta} = \eta$.

There is a limit, however, to how much the principal can adjust the transfers in this way: once t_L is made large enough, (IC_H) will start to bind and, although setting $q_L > q_H$ helps relax (IC_H), the fixed project size limits the principal's ability to provide incentives through implementation. Additionally, since contracting takes place *ex-ante*, the principal can induce truth-telling without surrendering a rent to the agent, so (PC) will bind as well.[12] Proposition 1 summarizes the results.

Proposition 1 (Costless Information Gathering): *When the agent is overconfident* $(\tilde{\eta} > \eta)$ *and information gathering is costless* $(c = 0)$*:*

- *Project implementation is efficient:* $q_L = 1$ *and* $q_H = 0$.
- *The principal's profit is strictly increasing in the degree of overconfidence and strictly larger than under agreement.*

Proof: All proofs for Section 2 can be found in Appendix A.

The two binding constraints uniquely determined the optimal transfers, so after substitution we have the principal's value function:

$$\Pi_0(\tilde{\eta}) = \eta[V - \beta_L] + (\tilde{\eta} - \eta)\Delta\beta$$

The first term, $\eta[V - \beta_L]$, is the maximum expected surplus from trade between the principal and agent when implementation is efficient, which the principal appropriates entirely due to contracting at the *ex-ante* stage. The second term, $(\tilde{\eta} - \eta)\Delta\beta$, measures the additional profit from

[12] The logic of the participation constraint being binding at the optimum carries through for the remainder of the paper.

the *participation effect* of overconfidence, which is positive and increase in the agent's belief. Notice that the benefits from the side-bet and from the agent's overvaluation of a given contract, each of which are captured in the participation effect, are distinct. To see this, suppose the principal offered some (t_L, t_H) and which set $t_L > t_H$, exactly satisfy (PC), and strictly satisfy both (ICs) under agreement. Allow the agent to become slightly overconfident and the principal can benefit from reducing both t_L and t_H proportionally if that proportion is small enough. This proportional reduction in the transfers does not constitute a side-bet (since the ratio of t_L to t_H remains unchanged), but benefits the principal, nevertheless. Hence, the *participation effect* measures the benefit the principal gets from the lower cost of inducing participation when the side-bet is written into the contract optimally.

1.2.b Costly Information Gathering

When the cost of information gathering effort is positive, the principal must now anticipate the agent's strategic choice of whether to become informed or not. If the optimal contract is to achieve efficient implementation, where $q_L = 1$ and $q_H = 0$, it will need to induce the agent to participate, gather information, and truthfully reveal what he has learned.[13] This requires introducing two *Information Gathering Incentive Compatibility* (IG) constraints, which are:

$$\tilde{\eta}[t_L - \beta_L] + (1 - \tilde{\eta})[t_H] - c \geq t_L - \beta_A \qquad (\boldsymbol{IG_L})$$

$$\tilde{\eta}[t_L - \beta_L] + (1 - \tilde{\eta})[t_H] - c \geq t_H \qquad (\boldsymbol{IG_H})$$

Satisfying these constraints ensures the agent will gather information about his type at cost $c > 0$ rather than claim to have learned β_L (resp. β_H) and implement (resp. shut down) the project.

[13] Any contract where $q_L = q_H$ need not induce information gathering, so $q_L = 1, q_H = 0$ is the only relevant case here.

With these constraints in hand, I will make the following assumptions:

$$c < \min\ \{\eta[V-\beta_L], -(1-\eta)[V-\beta_H]\} \qquad (\boldsymbol{A1})$$

$$c < \eta(1-\eta)\Delta\beta \qquad (\boldsymbol{A2})$$

Under (A1), gathering information is efficient under the principal's prior: the expected losses from implementing the project when implementation costs are high or failing to implement the project when implementation costs are low each exceed the cost of information gathering effort. Under (A2), inducing the agent's information gathering effort is free of agency costs under agreement ($\tilde{\eta} = \eta$): neither (IG) is binding and the principal can achieve her first-best outcome. Thus, any agency costs can be isolated and directly attributed to disagreement ($\tilde{\eta} > \eta$). Additionally, without (A2) the model would yield cases where overconfidence might either *encourage or discourage* information acquisition depending on the principal's prior and whether she believes gathering information is cost effective or not, but the former case has been studied before (see Goel and Thakor (2008), Gervais, Heaton and Odean (2011), Iossa and Martimort (2015)). Assumption (A2) allows me to emphasize a novel *disincentive* effect of overconfidence on information gathering. Finally, (A2) helps to highlight the fact that information may not be gathered in equilibrium due to the presence of *both* information asymmetry and overconfidence since under (A2) information acquisition is efficient in the absence of either one.

Including the information gathering constraints yields the following maximization problem for the principal, $[P_1]$:

$$\Pi_1(\tilde{\eta}) = \max_{t_H, t_L}\ \eta[V - t_L] + (1-\eta)[-t_H] \qquad \text{subject to}$$

$$\tilde{\eta}[t_L - \beta_L] + (1-\tilde{\eta})[t_H] - c \geq 0 \qquad (\boldsymbol{PC})$$

$$t_L - \beta_L \geq t_H \qquad (\boldsymbol{IC_L})$$

$$t_H \geq t_L - \beta_H \qquad (\boldsymbol{IC_H})$$

$$\tilde{\eta}[t_L - \beta_L] + (1 - \tilde{\eta})[t_H] - c \geq t_L - \beta_A \qquad (\boldsymbol{IG_L})$$

$$\tilde{\eta}[t_L - \beta_L] + (1 - \tilde{\eta})[t_H] - c \geq t_H \qquad (\boldsymbol{IG_H})$$

Observe that the (IGs) can be rewritten as:

$$\tilde{\eta}[t_L - \beta_L] \geq \tilde{\eta}(t_H) + c \qquad (\boldsymbol{IG_H})$$

$$(1 - \tilde{\eta})[t_H] \geq (1 - \tilde{\eta})[t_L - \beta_H] + c \qquad (\boldsymbol{IG_L})$$

which makes it clear that the (IGs) imply the (ICs). If the agent has an incentive to misreport in either state, he will be making the same report regardless of his information, and so information has no value to him. The contrapositive is that an agent who has learned his type will always truthfully reveal it.[14]

As in the case of costless information gathering, the principal has an incentive to write a side-bet into the contract by increasing t_L and reducing t_H, but in this situation the size of that bet is limited by (IG_L), rather than (IC_H), since such adjustments to the transfers make remaining uninformed and claiming $\beta = \beta_L$ more attractive (and claiming $\beta = \beta_H$ less attractive). These two binding constraints pin down the optimal transfers and, after substituting into (IG_H), yield the following condition for information gathering to feasible:

$$c \leq \tilde{\eta}(1 - \tilde{\eta})\Delta\beta \qquad (C1)$$

This is useful for characterizing the solution to problem $[P_1]$ and summarizing the results, as in Proposition 2.

Proposition 2 (Costly Information Gathering): *Suppose the principal wants to induce information gathering. When the agent is overconfident* $(\tilde{\eta} > \eta)$ *and information gathering is costly* $(c > 0)$, *there exits an* $\eta^* \in (\eta, 1)$ *such that:*

- *For all* $\tilde{\eta} \in (\eta, \eta^*]$, *(C1) is satisfied and:*

[14] This argument is analogous to that in Finkle (2005), where the principal gathers information rather than the agent.

- *Information gathering is feasible.*
- *Project Implementation is efficient:* $q_L = 1$ *and* $q_H = 0$.

- *For all* $\tilde{\eta} \in (\eta^*, 1)$, *(C1) is not satisfied and:*
  - *Information gathering is not feasible.*
  - *Project Implementation is not efficient:* $q_L = q_H$.

When the degree of overconfidence is not too high, $\tilde{\eta} \in (\eta, \eta^*]$, information gathering is feasible, and since the agent will truthfully report what he learns, the project is only implemented when costs are low. Thus, both information acquisition and implementation are efficient. However, overconfidence does increase the principal's cost of induce information gathering. The principal's value function is when (C1) holds is:

$$\Pi_1(\tilde{\eta}) = \eta[V - \beta_L] + (\tilde{\eta} - \eta)\Delta\beta - \frac{1-\eta}{1-\tilde{\eta}}c$$

This value function differs from that of the costless information gathering case by the term $-\frac{1-\eta}{1-\tilde{\eta}}c < -c$, which is negative and decreasing in the agent's overconfidence, and measures the *incentive effect*. The intuition behind this term is as follows: for a given contract that satisfies all the constraints, overconfidence increases the agent's payoff from becoming informed as well as his payoff from claiming $\beta = \beta_L$ while uninformed, but the benefit from the former effect is always entirely appropriated by the principal since an optimal contract extracts all *ex ante* rent from the agent (binds (PC)). The only effect that remains is the increased benefit of remaining uninformed, which causes (IG_L) to bind as well. In order to maintain the agent's indifference between gathering information and claiming $\beta = \beta_L$ while remaining uninformed, he must be increasingly rewarded for truthfully reporting $\beta = \beta_H$ to account for the increased benefit of remaining uninformed. This reward totals $\frac{c}{1-\tilde{\eta}}$, and is paid by the principal with probability $(1 - \eta)$. Thus, overconfidence increases the cost of inducing information gathering, and the principal's effort to exploit the agent at the contracting stage results in additional incentive costs at the information gathering stage.

When the degree of overconfidence is high enough, $\eta \in (\eta^*, 1)$, condition (C1) is violated and information gathering is not feasible. The agent's overconfidence makes it extremely attractive to forego gathering information and claim to have learned that $\beta = \beta_L$, since that report is only costly when it is wrong, and the agent expects to be wrong infrequently. To solve this problem, the principal can reward him for reporting that $\beta = \beta_H$, but that reward must be so large that the agent would prefer to stay uninformed and make that claim instead. In fact, any menu of transfers that satisfies one information gathering constraint will violate the other, so the agent cannot be motivated to gather information. Since inducing information gathering is not possible in this case, project implementation is necessarily inefficient: the optimal contract must involve implementing the project regardless of costs ($q_L = q_H = 1$) or never implementing the project at all ($q_L = q_H = 0$).

The infeasibility of information gathering is a stark example of how overconfidence can discourage learning, but it can also take a more subtle form. Consider, for a moment, an alternative model where the project's size can be varied continuously by the principal, i.e., $q \in [0, \infty)$, and the principal values the project according to an increasing and concave function $V(q)$. With more precise control over the project's size, the principal can use it to fine-tune the power of incentives for information gathering. This extra control will help in the sense that information gathering is feasible for any $\tilde{\eta} < 1$, but as the degree of overconfidence increases, providing the agent with sufficiently powerful incentives to gather information requires introducing large upward distortions in q_L and downward distortions in q_H. For low levels of overconfidence, these distortions only cut into the extra profit the principal gets from the participation effect. But once the agent is sufficiently overconfident, the distortions in the project's size become extreme, and the principal's profit will be decreasing in the degree of overconfidence.

1.2.c No Information Gathering

Whether information gathering is optimal or not when it is feasible depends on the principal's best alternative: a contract that does not induce information gathering. Such a contract can take two forms. Either the principal asks the agent to implement the project regardless of the costs in exchange for a fixed transfer, or the principal offers the "null" contract, where the project is shut down and the agent is paid nothing. Suppose the principals wants the project to be implemented in both states. Her problem is $[P_u]$:

$\Pi_u(\tilde{\eta}) = \max_{t_u} V - t_u$ subject to

$$t_u - \beta_A \geq 0 \qquad (\boldsymbol{PC})$$

$$t_u - \beta_A \geq \tilde{\eta}[t_u - \beta_L] + (1 - \tilde{\eta})[t_u - \beta_H] - c \qquad (\boldsymbol{NIG})$$

The solution to this problem is straightforward: any t_u strictly satisfies the *no information gathering constraint* (NIG), so the principal will set it to exactly compensate the agent for his subjective expected cost of implementation, $t_u = \beta_A$. The principal's value function is:

$$\Pi_u(\tilde{\eta}) = V - \beta_A$$

Since the principal is only concerned with the cost of production to the extent that she must compensate the agent for it, the agent's overconfidence, which reduces his expected cost of implementation, also increases the principal's profit. This represents another form of the participation effect, one in the absence of the side-bet that is written into the contract when the principal induces information gathering. For the remainder of the paper, I will assume that the principal's prior η is such that when information gathering is not induced, implementing the project in both states is optimal under that prior:

$$\Pi_u(\eta) = V - \beta_P > 0$$

It follows that this is also optimal under the agent's belief when he is overconfident, since $\beta_A < \beta_P$. This simplifies the comparison of contracts that do and do not induce information gathering and has no meaningful effect on the main results.

1.2.d When is Information Gathering Optimal?

It is now left to compare the principal's payoff from the information gathering contract to the no information gathering contract.

Proposition 3 (Optimal Information Gathering): *When the agent is overconfident and information gathering is costly:*

- *There exists an $\tilde{\eta}_x \in (\eta, 1)$ such that, for all $\tilde{\eta} \in (\tilde{\eta}_x, 1)$, the optimal contract does not induce information gathering, even when it is feasible.*
- *There exists a $\tilde{c} > 0$ and an $\tilde{\eta}_{max} \in (\eta, 1)$ such that for all $c \in (0, \tilde{c})$, it is the case that $\tilde{\eta}_{max} < \tilde{\eta}_x$, the principal's profit is strictly decreasing in $\tilde{\eta}$ for all $\tilde{\eta} \in (\tilde{\eta}_{max}, \tilde{\eta}_x)$ and increasing otherwise.*

The first part of the proposition captures the first main idea of the paper, that exploiting the differences in beliefs at the contracting stage may come at the cost of efficient information acquisition. A more overconfident agent is more willing to accept a given contract, yet less willing to gather information about his type. Though the principal may find information valuable, the cost of providing incentives gather it can outweigh those associated with inefficient project implementation.

The second part shows that increases in the degree of overconfidence may *harm* the principal on the margin. When the cost of information gathering is small enough, the principal's value function has a local maximum that is in the region where information gathering is both feasible and optimal. Thus, the principal's payoff is *non-monotonic* in the degree of overconfidence, increasing for low and high levels, but decreasing for intermediate levels. These regions correspond to (i) mild overconfidence, when information gathering is optimal, and the beneficial participation effect outweighs the costly incentive efficient, (ii) moderate overconfidence, when information gathering is optimal but now the incentive effect outweighs the participation effect on the margin, and (iii), extreme overconfidence, where information gathering is no longer optimal, and only the participation effect remains.

A similar non-monotonicity result can be found in Goel and Thakor (2008). In their model, a risk averse CEO (the agent) makes investment decisions on behalf of shareholders (the principal), and the agent is overconfident in the sense that he overestimates the precision of a signal he can obtain about investment returns. The agent's risk aversion makes his investment behavior too conservative from the perspective of the principal, but this conservatism is offset by mild overconfidence, which improves efficiency and increasing the principal's payoff. However, extreme overconfidence causes the agent to invest too aggressively, resulting in efficiency loss and reduction in the principal's payoff. Thus, the principal's payoff is non-monotonic in the degree of overconfidence, increasing when overconfidence is low, and decreasing when it is high.

One feature of their model that is crucial for that result is that the agent's degree of overconfidence is unknown to the principal, and even unknown to the agent himself at the time of contracting. Thus, there is no participation effect for the principal to exploit at the contracting stage, and any benefit the principal gains from overconfidence comes from the improved investment incentives. Further, if agent overconfidence were known to the principal, she could write contracts contingent on the degree of overconfidence that make her payoffs independent of it and neutralize any possible harm from overconfidence. My result complements theirs in showing that even when the degree of overconfidence is observable at the contracting stage, it may still cause problems for the principal. It also characterizes a new region, where the principal's payoff continues increasing in overconfidence once information gathering is no longer incentivized.

1.3 Diffidence

In the baseline model, I restrict attention to the case of overconfidence. The balance of economic theory concerned with differences in beliefs focuses on overconfidence, which suggests it is of more general interest. However, there is no reason *a priori* to rule out diffidence entirely, and it is not obvious whether the results in that case are analogous to that of overconfidence. This section shows that there is still an incentive effect, and it works against the principal as it does in the previous section.

Consider the model from Section 2 modified to reflect the agent's diffidence, i.e., $\tilde{\eta} \in (0, \eta]$. Assumptions (A1) and (A2) are maintained and serve the same role here. The principal's problem is $[P'_1]$:

$$\Pi'_1(\tilde{\eta}) = \max_{t_H, t_L, q_H, q_L} \eta[Vq_L - t_L] + (1-\eta)[Vq_H - t_H] \qquad \text{subject to}$$

$$\tilde{\eta}[t_L - \beta_L q_L] + (1-\tilde{\eta})[t_H - \beta_H q_H] - c \geq 0 \qquad (\boldsymbol{PC})$$

$$\tilde{\eta}[t_L - \beta_L q_L] + (1-\tilde{\eta})[t_H - \beta_H q_H] - c \geq t_L - \beta_A q_L \qquad (\boldsymbol{IG_L})$$

$$\tilde{\eta}[t_L - \beta_L q_L] + (1-\tilde{\eta})[t_H - \beta_H q_H] - c \geq t_H - \beta_A q_H \qquad (\boldsymbol{IG_H})$$

$$t_L - \beta_L q_L \geq t_H - \beta_L q_H \qquad (\boldsymbol{IC_L})$$

$$t_H - \beta_H q_H \geq t_L - \beta_H q_L \qquad (\boldsymbol{IC_H})$$

A few features of the model with overconfidence carry over. First, the participation constraint binds, as leaving a rent to the agent at the *ex-ante* stage remains suboptimal. Second, the information gathering constraints (IGs) imply the truth-telling constraints (ICs), so the (ICs) can once again be ignored. Lastly, the principal has the same motive to bet against the agent on the realization of his type. The main difference comes from the fact that the side-bet will now involve an increase in t_H and decrease in. Naturally, increasing t_H and decreasing t_L makes remaining uninformed and claiming $\beta = \beta_H$ more attractive to the agent, so (IG_H) binds at the optimum, while (IG_L) may or may not. Substituting the binding constraints into the objective and (IG_L) yields the simplified problem $[P_1']$:

$$\Pi'_1(\tilde{\eta}) = \max_{q_H, q_L} \eta[(V - \beta_L)q_L] + (1-\eta)[(V - \beta_H)q_H]$$

$$-\frac{\eta}{\tilde{\eta}}c - (\eta - \tilde{\eta})\Delta\beta q_H \qquad \text{subject to}$$

$$c \leq \tilde{\eta}(1-\tilde{\eta})\Delta\beta(q_L - q_H) \qquad (IG_L)$$

Observing that increasing q_L and decreasing q_H both increase the objective and relax (IG_L), it is optimal to set $q_L = 1$ and $q_H = 0$. Plugging these values into (IG_L) yields the condition (C1) from the baseline model:

$$c \leq \tilde{\eta}(1-\tilde{\eta})\Delta\beta \qquad (C1)$$

I can now state the analogue to Proposition 2 for the case of diffidence.

Proposition 2.A.*: Suppose the principal wants to induce information gathering. When the agent is diffident* $(\tilde{\eta} < \eta)$ *and information gathering is costly* $(c > 0)$*, there exists an* $\eta^* \in (0, \eta)$ *such that:*

- *For all* $\tilde{\eta} \in [\eta^*, \eta)$*, (C1) is satisfied and:*
    - *Information gathering is feasible.*

- *Project Implementation is efficient:* $q_L = 1$ *and* $q_H = 0$.

- *For all* $\tilde{\eta} \in (0, \eta^*)$, *(C1) is not satisfied and:*
  - *Information gathering is not feasible.*
  - *Project Implementation is not efficient:* $q_L = q_H$.

Proof: All Proofs for Section 3 can be found in Appendix B.

For mild diffidence (C1) is satisfied, information gathering is feasible, and the principal's value function is:

$$\Pi'_1(\tilde{\eta}) = \eta[V - \beta_L] - \frac{\eta}{\tilde{\eta}}c$$

It is useful to compare to the value function from the case of overconfidence:

$$\Pi_1(\tilde{\eta}) = \eta[V - \beta_L] + (\tilde{\eta} - \eta)\Delta\beta - \frac{1-\eta}{1-\tilde{\eta}}c$$

The term associated with the expected surplus from implementation, $\eta[V - \beta_L]$, is unchanged. While there is a term associated with the participation effect in the case of overconfidence, $(\tilde{\eta} - \eta)\Delta\beta$, there is no such term in the case of diffidence when $q_H = 0$.[15] When $q_H = 0$, the benefit from the side-bet written into the transfers exactly offsets the cost of the higher expected transfer needed to induce participation. The new incentive effect is given by $-\frac{\eta}{\tilde{\eta}}c < -c$, which is negative and *decreasing* as the agent becomes more diffident ($\tilde{\eta}$ decreases). The intuition is like that of the case of overconfidence: misreporting his cost as $\beta = \beta_H$ when he is uninformed is only costly to the agent when his realized cost turns out to be $\beta = \beta_L$, which he believes is increasingly unlikely when he becomes increasingly diffident. Thus, increasing diffidence makes this deviation more attractive and deterring it with a reward (of $\frac{c}{\tilde{\eta}}$) when the agent reports $\beta = \beta_L$ more expensive.

---

[15] Suppose the project size remained binary but always strictly positive: $q \in \{1, \underline{q}\}$ where $1 > \underline{q} > 0$. Since $q_H = \underline{q}$ at the optimum, the participation effect with diffidence would be given by $-(\eta - \tilde{\eta})\Delta\beta\underline{q} < 0$.

For extreme diffidence ($\tilde{\eta} < \eta^*$), (C1) is violated and the results are the same as with overconfidence. Information gathering is not feasible as rewards meant to deter remaining uniformed and claiming $\beta = \beta_H$ must be so large that instead claiming $\beta = \beta_L$ becomes a profitable deviation for the agent, and vice versa.

Once again, the relevant alternative is a contract that does not induce information gathering. Although the analysis of the contract that requests that the agent implement the project without knowledge of his costs is unchanged from Section 2.c., care must be taken now as the value function from that contract, $\Pi_u(\tilde{\eta}) = V - \beta_A$, is positive for $\tilde{\eta}$ large enough but necessarily falls below zero as $\tilde{\eta} \to 0$. The following proposition summarizes.

Proposition 4: *Suppose the principal does not want the agent to gather information. Let* $\eta^{**} = -\frac{V-\beta_H}{\Delta\beta} > 0$. *When the agent is diffident* $(\tilde{\eta} < \eta)$ *and information gather is costly* $(c > 0)$*:*

- *For all* $\tilde{\eta} \in (\eta^{**}, \eta)$, *the optimal contract is such that* $q_L = q_H = 1$ *and the principal's profit is* $\Pi_u(\tilde{\eta}) = V - \beta_A > 0$.
- *For all* $\tilde{\eta} \in (0, \eta^{**}]$, *the optimal contract is the null contract,* $q_L = q_H = 0$ *and* $t_L = t_H = 0$*, and the principal's profit is* $\Pi_u(\tilde{\eta}) = 0$.

It is now possible to state the analogue to Proposition 3 for the case of diffidence.

Proposition 3.A. (Optimal Information Gathering): *When the agent is diffident and information gathering is costly:*

- *There exists an* $\tilde{\eta}_x \in (0, \eta)$ *such that for all* $\tilde{\eta} \in (0, \tilde{\eta}_x)$, *information gathering is not induce in the optimal contract, even when it is feasible.*
- *The principal's profit is (weakly) increasing in* $\tilde{\eta}$ *for all* $\tilde{\eta} \in (0, \eta)$.

The first part of the proposition describes the optimality of inducing information gathering. When the agent is diffident enough, even if information gathering is feasible it will be suboptimal;

it is the mirror image of the first part of Proposition 3. The second part demonstrates the simple intuition that is captured in many previous studies of non-common priors in contracting: increasing diffidence reduces the principal's payoff.

Why is it that excessive diffidence has the same effect on incentives to gather information as excessive overconfidence? Once again, the issue is the principal's motive to bet against the agent and the effect it has on the agent's information gathering incentives. By writing a side-bet into the contract, the principal causes one of the information gathering constraints to bind, making the agent indifferent between gathering information and making one uninformed claim about his costs or another. However, that claim is only costly to the agent when it turns out to be wrong *ex post*. Diffidence causes the agent to underestimate the probability that he is wrong when announcing β_H in the same way that overconfidence causes the agent to underestimate the probability that he is wrong when announcing β_L.

This argument makes it clear that there is a relationship between the *precision* of the agent's belief (which increases as $\tilde{\eta}(1-\tilde{\eta}) \to 0$) and his willingness to gather information.[16] As is mentioned in the introduction, Goel and Thakor (2008) and Gervais, Heaton and Odean (2011) each consider information gathering models where overconfidence is modeled explicitly as over precision, but what is meant by over precision in their work differs from what it means here. In each of those papers, the agent believes the signal he gets from gathering information is more precise than it is; over precision in this sense unambiguously increases the value of gathering more

[16] In an earlier version of this paper (Downs (2019)), I included an extension to the main model that involved a continuous choice of project size ($q \in [0, \infty)$ and a continuous type space ($\beta \in [\beta_L, \beta_H]$ where $0 < \beta_L < \beta_H$). This extension allows for the definition of *overly precise* beliefs while holding the mean of those beliefs fixed. As one might expect, the participation effect disappears when the principal's and agent's beliefs share a mean, but the negative incentive effect is present when the agent's belief is a mean preserving contraction of the principal's.

information. In this paper, the agent's *prior* is overly precise when he is sufficiently overconfident or diffident, which unambiguously *decreases* the value of gathering more information.

1.4 Conclusion

In this paper I have studied the effects of overconfidence on information acquisition in a principal-agent relationship. I have shown that overconfidence can reduce an agent's incentive to learn and that increases in overconfidence are not necessarily beneficial to the principal. These results contrast with existing literature where overconfidence is generally predicted to encourage information acquisition and increase the principal's payoff.

This model lends itself to several interesting extensions.[17] As is typical in the information gathering literature, it is assumed that the principal is precluded from gathering information about the agent's cost herself. The results suggest that the principal would prefer to do so if she could report what she has learned to the agent credibly. If not, this would introduce a new incentive cost in that the contract would need to be designed to convince the agent that any report made by the principal is truthful. Finkle (2005) studies information gathering by a principal but does not consider the possibility that the principal and agent disagree on the distribution of types.

Another natural extension is to allow the value of the agent's outside option to vary with his overconfidence. One might expect that the agent's overconfidence about his cost of implementation would also extend to other activities or professions. If the agent's outside option becomes more attractive in his degree of overconfidence, the participation effect of overconfidence would be attenuated, which cuts in the same direction as the second part of Proposition 3, that increases in overconfidence do not necessarily help the principal. However, Van Den Steen (2004)

[17] The main results are robust to changes in the timing of the game as well as the agent's risk preference. See Downs (2019).

argues that an agent who chooses tasks according to which is most likely to succeed is more likely to choose a task for which he overestimates his chance of success. An agent's outside option may very well be an alternative he expects to fail, causing him to underestimate his expected payoff from that alternative. If this is the case, the participation effect will be exaggerated, which may attenuate or overturn some of this paper's main results.

Finally, it has been assumed throughout the paper that the players' beliefs are common knowledge.[18] This assumption is restrictive; if the principal were to offer the contracts described in Propositions 2 and 2.A. as a menu to agents with private information about their beliefs, at least some agents would have an incentive to understate their level of confidence. The principal will have to strike a balance between surrendering rent to the most overconfident agents and offering the null contract to some mildly diffident agents that would otherwise produce, which helps reduce those rents. The precise qualitative features of such a screening contract will depend on the distribution of beliefs the principal expects to face. I leave this and other extensions mentioned above for future research.

[18] See, for example, Eliaz and Spiegler (2008) for a contracting model where the agent's beliefs are his private information.

Chapter 2

Overconfidence, Competition, and Market Efficiency

Abstract: In a labor market with moral hazard, adverse selection and overconfident workers, I compare the equilibrium contracts and efficiency of perfectly competitive markets and monopsonistic markets. Screening workers according to their productivity (type) requires both competitive firms and a monopsonist to distort the contract offered to less productive workers, and these distortions may offset others that arise due to overconfidence. In each market structure, the overall effect on efficiency depends on the type-distribution of the labor force and/or the relative degree of overconfidence of the different types of the worker. I provide sufficient conditions for a monopsonistic market to be more efficient than a competitive market, but also that overconfident workers prefer to be employed in competitive markets, where their overconfidence harms them the most.

2.1 Introduction

In a wide variety of contexts, individuals tend to overestimate their own abilities or prospects for the future. For example, entrepreneurs overestimate their ability to generate a profit in competitive markets (Camerer and Lovallo 1999), CEOs overestimate their ability to identify and manage value-enhancing mergers (Malmendier and Tate 2008), and workers, especially men, tend to overestimate their productivity within a firm (Sautmann 2013).[19] Firms that interact with these individuals on a regular basis are likely to be aware the individual's, and contracts between the parties will reflect that. This can result in contracts that "exploit" the individual's overconfidence, even in perfectly competitive markets, where one might suspect that competition would discipline firms against exploitative behavior. When equilibrium contracts in perfectly competitive markets entail such exploitative inefficiencies, are they necessarily more efficient than less competitive markets?

To answer this question, I consider a labor market with moral hazard, adverse selection, and workers who are overconfident in the sense that they overestimate their productivity. In this market, firms hire workers to exert effort producing a good. Neither the worker's effort nor his productivity is observable, so firms offer labor contracts consisting of a fixed base salary and an output-contingent bonus meant to align the incentives of the worker for efficient effort with that of the firm. Workers who overestimate their productivity also overestimate the returns to their effort, which gives them a preference for highly variable compensation packages which involve low base salaries and the potential for large bonuses, i.e., high-powered incentive contracts.

If each worker's productivity were observable, both competitive labor markets, where free entry leads to employers earning zero profit, and monopsonistic labor markets, where a single employer

[19] See also, Barber and Odean (2000), Malmendier and Tate (2005), Santos-Pinto (2008), DellaVigna (2009), Larkin and Leider (2012), Ben-David et al (2013), Pikulina et all (2017), He et al (2019), or Aktas et al (2019).

earns positive profit, would respond to the preferences of workers by offering incentive contracts which induce inefficiently high effort. These upward distortions in the workers' effort come about for different reasons in different market structures. In competitive markets, firms bid for the services of workers, and competitive pressure forces them to offer the high-powered incentive contracts that workers most prefer. Since these distortions are the result of firms catering to the preferences of workers, I refer to it as the *enablement* of workers.[20] The monopsonist, on the other hand, requests high effort levels in order to take advantage of the errors workers make in forecasting their payoff. High-powered incentives induce inefficiently high effort by workers in equilibrium, which exacerbates the workers' overestimation of their payoff and allows the monopsonist to reduce the fixed portion of the worker's compensation, to her benefit. Since the inefficiently high effort levels in monopsony come about due to the monosponist's profit motive, and at the expense of workers, I refer to it as the *exploitation* of the workers' overconfidence.

When the workers' productivities (types) are not observable, firms must now concern themselves with screening workers since more productive workers (high types) may have an incentive to misrepresent themselves as less productive (low types). Qualitatively, competitive firms and a monopsonist separate types in a similar way, by offering relatively low-powered incentive contracts to low types, which are relatively less attractive to high types. There are, however, important quantitative differences which are a consequence of the differing goals firms have in the different market structures.

In monopsony, the firm's problem exhibits features that are similar to a standard screening model without overconfidence: high types are offered contracts which induce the same effort level as if the types were observable, sometimes referred to as "no distortion at the top", while low types

[20] Firms earn zero profit in any competitive equilibrium, so their payoff is invariant to the biases of workers.

are offered contracts that distort their effort downward, which helps extract information rent from high types. With overconfident workers, high-type effort is unambiguously above the efficient level, to the point where the marginal cost of this inefficiency is exactly offset by the marginal benefit to the principal of exploiting the worker's overconfidence. On the other hand, low-type effort may be at, above, or below the efficient level, depending on the intensity of the trade-off between extracting rent from high types, exploiting low types, and maximizing the surplus generated by low types. Importantly, it is the monopsonist's profit motive that leads to rent extraction, and downward distortions of low-type effort due to rent extraction can partially or entirely offset the upward distortions due to overconfidence.

In competitive markets, firms may or may not need to screen types. When the (true) productivity difference between high types and low types is small enough, contracts specify effort levels that are efficient in the eyes of the workers, i.e., which maximally *enable* the workers' overconfidence, and which are incentive compatible. Since worker overconfidence leads to a preference for high-powered incentives, the equilibrium effort levels of both types are unambiguously above the efficient levels. When the productivity difference between types is large enough, competitive firms must distort the effort of low types downward in order to achieve incentive compatibility. The key difference between this distortion and the one introduced by a monopsonist is that, with no profit motive, the competitive firm reduces the effort of low types to the minimum extent possible to keep the contract as attractive to low types as possible, a result of competitive pressures this firm faces from others. This does *partially* but *never completely* offsets the upward distortion due to overconfidence, and so low-type effort remains unambiguously above the efficient level.

There are immediate welfare implications. First, regardless of the degree of competition, distortions due to screening and overconfidence can interact to improve overall market efficiency relative to markets where one or the other is absent. Thus, neither correcting overconfident beliefs in the presence of asymmetric information, nor reducing asymmetric information in the presence of overconfidence will necessarily improve market outcomes. Second, labor markets where employers have some market power may prove to be more efficient than those where firms have none, as market power yields positive profits, which gives firms an (indirect) incentive to mitigate inefficiencies that are driven by overconfidence. Lastly, overconfident workers prefer more competitive markets, even when they are less efficient. This effect is primarily a result of the increased bargaining power workers enjoy in competitive markets, but has a second component, which is that the high-powered incentive contracts that overconfident workers prefer tend to prevail in more competitive markets.[21]

This paper contributes to the literature on contracting with overconfident agents. One common theme in this literature is that contracts offered by principals to overconfident agents will exploit that overconfidence to the benefit of the principal. This can happen when an agent overestimates his own ability,[22] mispredicts his future behavior,[23] or underestimates the risk he faces.[24] Some papers have also shown that perfect competition does not necessarily prevent inefficiencies from arising due to behavioral biases,[25] and others that overconfidence and asymmetric information can cause mutually attenuating effects, giving rise to contracts that are

[21] In Section 3, I show that less productive and overconfident workers would be willing to pay a positive cost to credibly communicate their type to competitive firms. This result does not hold in the absence of overconfidence.
[22] For example, if he overestimates his productivity or underestimates his cost of production, both of which are considered herein.
[23] See DellaVigna and Malmendier (2006), Eliaz and Spiegler (2008), Grubb (2009), or Grubb and Osborne (2015).
[24] See Goel and Thakor (2008) or Gervais, Heaton, and Odean (2011).
[25] See Manove and Padilla (1999), Gabaix and Laibson (2006), Heidhues and Koszegi (2010) or Grubb (2015). See also Heidhues and Koszegi (2015) for several more relevant references.

more efficient than in the absence of either one.[26] This paper reproduces all of those features; exploitative contracts, inefficient competitive markets, and offsetting effects of overconfidence and asymmetric information, all in a simple labor market model.

There are only a few papers that consider both competitive and monopsonistic (or monopolistic) screening in the presence of overconfidence. In Grubb (2009), buyers in the market for cell phone plans underestimate the variance of their demand, to which firms optimally respond by offering exploitative three-part tariffs. This happens in both monopolistic and competitive markets, resulting in inefficiencies in either case. Sandroni and Squintani (2013) study insurance markets with both low-risk and high-risk types, and where some high-risk types believe they are low risk. They derive the equilibrium insurance contracts in both monopolistic and perfectly competitive markets, and show that the presence of overconfidence may be observationally equivalent to a change in the type distribution in monopoly, but that it results in qualitative changes in the equilibrium contracts under perfect competition. While each of these papers consider the effect of overconfidence on equilibrium contracts in both monopoly and perfect competition, neither present welfare comparisons between the two. This paper's main contribution is the explicit welfare comparisons between monopsony and perfect competition in the presence of overconfidence and asymmetric information. These analytical welfare comparisons make it possible to do comparative statics, to analyze the effect of policy on each type of market and can inform empirical strategies for understanding the role of overconfidence in markets with overconfident decision-makers. I am able to obtain these analytical results by taking advantage of

[26] See Gervais and Goldstein (2007), Goel and Thakor (2008), or De La Rosa (2011).

the simple and tractable framework of Benabou and Tirole (2016) (BT hereafter), who also study the interaction between screening distortions, competition, and social welfare.[27]

The paper is organized as follows. The baseline model is introduced in Section 2, along with analysis of equilibrium contracts in both competitive and monopsonistic markets. Welfare analysis for both the overall market and individual workers is in Section 3. In Section 4, I consider the possibility that workers may be under-confident and provide conditions under which the main results hold, as well as an alternative model that supports the robustness of the main results. Section 5 concludes.

2.2 Model

Firms hire workers to produce a good or implement a project. Gross revenue collected by the firm is given by $A \cdot y = A \cdot (\theta + e)$, where $A > 0$ represents the quality of the firm's technology or intensity of demand for the good, $y = \theta + e$ is the output produced by the worker or the quality of the implemented project, $\theta \geq 0$ is the worker's productivity and type, and $e \geq 0$ is the worker's effort.[28] While firms do not observe either e or θ directly, the worker's total contribution to gross revenue y is observable, verifiable, and contractible. Firms compensate workers for the effort, which has a disutility of $c(e) = \frac{1}{2}e^2$, with a linear compensation scheme (β, α), where β represents a piece rate or bonus rate and α represents a fixed wage/fee or salary. A firm's profit is given by:

---

[27] Benabou and Tirole (2016) study a labor market with adverse selection and moral hazard in multiple tasks. One task is measured more precisely than the other, so monetary incentives are optimally tied to the more precisely measured task. Increased competition leads to over-incentivization of the measured task, which crowds out effort on the unmeasured task, leading to efficiency losses. They show that social welfare is maximized at a strictly intermediate degree of competition between monopsony and perfect competition. More on how my results compare to theirs in Section 3.

[28] The worker's output, $y = \theta + e$, takes a linear form here for simplicity, but all the main results can be obtained in a model with the multiplicative form $y = \theta e$. See Section 4.2.

$$\Pi(\alpha,\beta;e,\theta) = Ay - \beta y - \alpha = (A-\beta)(\theta+e) - \alpha$$

and a worker's payoff is given by:

$$U(e;\alpha,\beta,\theta) = \beta y + \alpha - \frac{1}{2}e^2 = \beta(\theta+e) + \alpha - \frac{1}{2}e^2.$$

Both workers and firms have an outside option with payoff normalized to zero.[29]

The relationship suffers from two types of informational asymmetry: the worker's effort e is not observed by firms and the worker is privately informed about his productivity θ. Workers are *overconfident* in the sense that they overestimate their productivity within the firm.[30] That is, firms believe that $\theta \in \{\theta_L, \theta_H\}$, where $\theta_H > \theta_L > 0$ and $0\Delta\theta \equiv \theta_H - \theta_L > 0$, while workers believe that $\theta \in \{\hat{\theta}_L, \hat{\theta}_H\}$ where $\hat{\theta}_H > \hat{\theta}_L > 0$ and $\Delta\hat{\theta} \equiv \hat{\theta}_H - \hat{\theta}_L > 0$. Let:

$$\hat{\theta}_i \equiv \theta_i + \gamma_i$$

for $i \in \{L, H\}$. The parameter γ_i measures the difference in beliefs between firms and workers and, to restrict attention to overconfident workers, I assume that $\gamma_i > 0$ for $i \in \{L, H\}$.[31] These beliefs about θ are all common knowledge, and firms and workers *agree to disagree* about the true values of θ.[32] Finally, while firms and workers disagree on how productive a "low type" or a "high type" worker is, they agree on the distribution of types, where q is the probability that the worker is a low type and $1-q$ is the probability that the worker is a high type.[33]

---

[29] The assumption that the worker's outside option yields him zero utility is for analytical convenience but results in $\alpha < 0$ in equilibrium. Instead, it may be assumed that the worker's outside option is worth some $\overline{U} > 0$ large enough that $\alpha > 0$ in equilibrium, but not so large that firms exclude low types. A non-empty range of $\overline{U}$ satisfying these conditions exists as long as A is large enough, and all the main results are robust to this alteration. A proof along these lines can be found in BT, Online Appendix D.

[30] Hence the type-invariant and overconfidence-invariant outside option. This may be the case when workers gravitate towards labor markets where they perceive themselves to be most able.

[31] I allow for diffidence (under-confidence) in Section 4.

[32] This is a common assumption in models of contracting with non-common priors that avoids contracts that send signals about the principal's (firm's) beliefs.

[33] Since contract takes place at the interim stage, i.e., after the worker has privately observed his own type, disagreement on the distribution of types is of no consequence: no decision the worker makes depends on that distribution. So, if firms believed that the proportion of low types was $\hat{q} \neq q$, all the main results hold with all q replaced with a corresponding $\hat{q}$.

The timing of the game is as follows:

$t = 0$: Each worker forms a belief about his type, i.e. privately observes $\theta \in \{\hat{\theta}_L, \hat{\theta}_H\}$.

$t = 1$: Firms offer a menu of contracts $(\beta_i, \alpha_i)_{i\in\{L,H\}}$ to the workers.

$t = 2$: Firms and the worker exchange the fixed payment α_i.

$t = 3$: The worker privately exerts effort e_i.

$t = 4$: Output y is realized and publicly observed. The firm collects $A(\theta_i + e_i)$ and pays the worker $\beta_i(\theta_i + e_i)$. Payoffs are realized and the game ends.

2.2.1 Worker Effort and Market Efficiency

After accepting the contract (β, α), the worker chooses his effort by solving the following maximization problem:

$$\max_{e} U(e; \alpha, \beta, \theta) = \max_{e} \beta(\theta + e) + \alpha - \frac{1}{2}e^2.$$

The solution to this problem sets $e = \beta$. Since this relationship must hold in equilibrium, for the remainder of Sections 2 and 3, I will replace the bonus rate β with the effort level e, and discuss effort levels directly. Thus, a menu of contracts can be written as $(e_i, \alpha_i)_{i\in\{L,H\}}$, the worker's indirect utility from accepting a contract (e, α) is:

$$U(e; \alpha, \theta) = e\theta + \alpha - \frac{1}{2}e^2$$

and a firm's payoff is:

$\Pi(\alpha, e; \theta) = (A - e)(\theta + e) - \alpha.$

In order to uncover the efficiency and welfare properties of markets with asymmetric information, overconfidence workers, and varying degrees of competition, I will assume explicitly here that it is the *firms'* beliefs about θ that are the true, objective values, and maintain that assumption for all of Sections 2 and 3.[34] Then, define social surplus when the worker's type is i as:

$$S_i(e_i) \equiv A(\theta_i + e_i) - \frac{1}{2}(e_i)^2$$

which represents the objective value of the project minus its cost, given effort level e. This measure of welfare coincides with the sum of the worker's *experienced* (*ex post*) utility and profit of the firm he workers for (which is the same *ex ante* and *ex post*), so it can be understood as measuring *ex post* welfare.[35] From this definition, it is clear that social surplus is maximized at $e_i^* = A$, and the maximized surplus is:

$S_i^* \equiv S_i(e_i^*) = \frac{1}{2}A^2 + A\theta_i$.

I will refer to $e^* = A$ as the *efficient level of effort*, which serves as a natural point of comparison when studying optimal contracts in the presence of asymmetric information, overconfidence, or both. Finally, I will assume that A is large enough that $S_i(e_i)$ is strictly positive in any equilibrium.

2.3 Monopsony

2.3.1 Observable type

Consider the case of a single firm (the monopsonist, she) offering contracts to workers as in a standard monopolistic screening model. Suppose for the moment, that the firm is able to observe the worker's type at the time of contracting.[36] The firm solves the following problem for each type of worker $i \in \{L, H\}$:

---

[34] This assumption, which is also used in Grubb (2009), is meant to capture the idea that while individuals may tend to evaluate themselves too favorably, firms with experience contracting with such individuals will be aware of this bias, even if they cannot observe the worker's type directly.

[35] See Kahneman et al (1997) for a discussion of the use of experienced utility as a measure of welfare.

[36] That is, the firm observes that the worker is type θ_i and knows the worker's belief is that his type is $\hat{\theta}_i$.

$$P_i^{MO}: \max_{e_i, \alpha_i} (A - e_i)(\theta_i + e_i) - \alpha_i$$

$$[IR_i]: e_i\left(\hat{\theta}_i + e_i\right) + \alpha_i - \frac{1}{2} e_i^2 \geq 0$$

The superscript MO indicates monopsony when types are observable. Though the solution to this problem is straightforward, it is worth noting that, after substituting the binding individual rationality constraints $[IR_i]$ and using the definition of social welfare given in the previous section, the problem can be restated as:

$$P_i^{MO}: \max_{e_i} S_i(e_i) + \gamma_i e_i$$

In the absence of overconfidence ($\gamma_i = 0$), the monopsonist would offer contracts that maximize social surplus, and appropriate it entirely through the fixed fees α_i. She can do better, however, with overconfidence workers ($\gamma_i > 0$) by offering contracts that induce a higher effort level and charging the worker a fixed fee that is actually large than the maximized social surplus. This is evident in the solution to problem P^{MO}:

$$e_i^{MO} = A + \gamma_i$$

$$\alpha_i^{MO} = -\left[S_i^* + \frac{3}{2}\gamma_i^2 + 2A\gamma_i + \theta_i\gamma_i\right]$$

This contract illustrates a key feature of the monopsony case: the worker's overconfidence allows him to be exploited by contracts that induce effort above the efficient level. Since he is convinced that his effort is more valuable that it is, he also overestimates the size of his bonus for a given effort level. The fixed payment he makes to the firm is set larger than the surplus he actually generates, leaving him worse off than had he taken his outside option, and the firm better off than under agreement.[37]

[37] Agreement being the case where the workers hold the same (unbiased) beliefs about their productivity as firms.

2.3.2 Unobservable Type

Now, suppose that the worker's type is not observed by the firm before contracting. The firm's problem is:

$$P^M: \max_{e_i, \alpha_i} q[(A - e_L)(\theta_L + e_L) - \alpha_L] + (1 - q)[(A - e_H)(\theta_H + e_H) - \alpha_H]$$

$$[IR_L]: e_L\left(\hat{\theta}_L + e_L\right) + \alpha_L - \frac{1}{2}e_L^2 \geq 0$$

$$[IR_H]: e_H\left(\hat{\theta}_H + e_H\right) + \alpha_H - \frac{1}{2}e_H^2 \geq 0$$

$$[IC_L]: e_L\left(\hat{\theta}_L + e_L\right) + \alpha_L - \frac{1}{2}e_L^2 \geq e_H\left(\hat{\theta}_L + e_H\right) + \alpha_H - \frac{1}{2}e_H^2$$

$$[IC_H]: e_H\left(\hat{\theta}_H + e_H\right) + \alpha_H - \frac{1}{2}e_H^2 \geq e_L\left(\hat{\theta}_H + e_L\right) + \alpha_L - \frac{1}{2}e_L^2$$

The equilibrium contracts are as follows.

Proposition 1 (Monopsony): *The solution to P^M is:*

$$e_H^M = A + \gamma_H$$

$$e_L^M = A + \gamma_L - \frac{1-q}{q}\Delta\hat{\theta}$$

$$\alpha_H^M = \alpha_H^{MO} + \left[A + \gamma_L - \frac{1-q}{q}\Delta\hat{\theta}\right]\Delta\hat{\theta}$$

$$\alpha_L^M = \alpha_L^{MO} + \frac{1-q}{q}\Delta\hat{\theta}\left[A + 2\gamma_L + \theta_L - \frac{1-q}{2q}\Delta\hat{\theta}\right]$$

High-type effort is above the efficient level, and low-type effort is above the efficient level if and only if $q > \frac{\Delta\theta + \Delta\gamma}{\Delta\theta + \gamma_H} \equiv q_L \in (0,1)$.

All proofs for Sections 2 and 3 can be found in Appendix A.

Problem P^M is a typical monopolistic screening problem, and, since $\Delta\hat{\theta} > 0$, the solution also has typical properties that $[IR_L]$ and $[IC_H]$ bind, while the other constraints are slack. After substituting in the binding constraints, the objective of problem P^M becomes:

$$\max_{e_i} q[S_L(e_L) + \gamma_L e_L] + (1-q)[S_H(e_H) + \gamma_H e_H] - (1-q)\Delta\hat{\theta} e_L$$

Three terms motivate the monopsonist to deviate from the efficient effort levels. The first two, $\gamma_L e_L$ and $\gamma_H e_H$, are present for the same reason as when the worker's type was observable: the principal requests effort levels above the efficient level in order to exploit the worker's bias. The last one, $-(1-q)\Delta\hat{\theta} e_L$, corresponds to information rent surrendered to high-type workers, which can be reduced by reducing e_L. While the high type's effort is unambiguously above the efficient level, there are now offsetting distortions in the effort level of low types, upward due to overconfidence and downward due to rent extraction. When q is large enough, the overall distortion is still upward, but less so than when the worker's type is observable, meaning the introduction of asymmetric information has made the market unambiguously more efficient.[38]

2.4 Perfect Competition

2.4.1 Observable Type

Consider the case of a labor market where many firms compete for the services of workers, and where entry into the market is costless. When the worker's type is observable, firms solve the following problem for each type of worker $i \in \{L, H\}$:

$$P_i^{CO}: \max_{e_i, \alpha_i} \left[e_i(\hat{\theta}_i + e_i) + \alpha_i - \frac{1}{2e_i^2} \right]$$

$$[ZP_i]: (A - e_i)(\theta_i + e_i) - \alpha_i = 0$$

[38] More on this in Proposition 4.

The superscript CO indicates perfectly competitive markets with observable types. Firms competing for workers offer contracts that maximize the worker's utility among all contracts that earn non-negative profit, and free entry pins down equilibrium profits at exactly zero. Substituting the zero-profit condition into the objective yields and using the definition of social surplus, the problem simplifies to:

$$P_i^{CO}: \max_{e_i} S_i(e_i) + \gamma_i e_i$$

The objective is identical to the monopsony case, and so the same effort levels are optimal, while the fixed payments can be obtained from the zero-profit conditions:

$$e_i^{CO} = A + \gamma_i$$

$$\alpha_i^{C0} = -\gamma_i(A + \theta_i + \gamma_i)$$

While the effort levels are the same as in monopsony with observable types, the warrant a new interpretation. Like in monopsony, a worker's overconfidence causes him to overestimate the value of his effort, which results in contracts that specify inefficiently high levels of effort. However, firms do not benefit from this distortion at all (apart from attracting workers to begin with) since they earn zero profit in equilibrium, so it is not appropriate to refer to this as exploitation. Rather, firms offer these contracts to *enable* the workers: overconfidence gives the worker's a preference for higher levels of effort, and firms cater to this to attract them. The fixed payments, which would be zero in the absence of overconfidence, are once again negative, but enough to maintain zero profit for firms.[39]

[39] For a point of comparison, Sandroni and Squintani (2007,2013) each study a perfectly competitive insurance market where overconfidence is modeled as some high-risk consumers believing they are low risk. High-risk consumers are always fully insured, while low-risk and overconfident consumers are underinsured to varying degrees, depending on the number of overconfident consumers. If consumers in their model were overconfidence in a sense analogous to mine, where each type underestimates their probability of an accident, both types of consumers would receive less than full insurance, and screening distortions would exacerbate the underinsurance for low-risk consumers.

2.4.2. Unobservable Types

For the case of competitive markets with unobservable types, I will follow the logic of BT Section 3.B. First, I will solve for a menu of incentive-compatible contracts that yield a firm offering that menu zero profit *for each type of worker* (i.e., that does not involve cross-subsidization between types). This is the solution concept used in Rothschild and Stiglitz (1976), and so I will refer to these contracts as the *RS* contracts. Second, I will derive conditions under which the RS contracts are the unique competitive equilibrium contracts.

When the worker's type is not observed, firms still seek to offer a menu of contracts that yield their workforce the highest utility possible but need that menu to be incentive-compatible in order to maintain non-negative profit. These contracts can be characterized as the solution to the following problem:

$$P^{RS}: \max q\left[e_L(\hat{\theta}_L + e_L) + \alpha_L - \frac{1}{2}e_L^2\right] + (1-q)\left[e_H(\hat{\theta}_H + e_H) + \alpha_H - \frac{1}{2}e_H^2\right]$$

$$[ZP_L]: (A - e_L)(\theta_L + e_L) - \alpha_L = 0$$

$$[ZP_H]: (A - e_H)(\theta_H + e_H) - \alpha_H = 0$$

$$[IC_L]: e_L(\hat{\theta}_L + e_L) + \alpha_L - \frac{1}{2}e_L^2 \geq e_H(\hat{\theta}_L + e_H) + \alpha_H - \frac{1}{2}e_H^2$$

$$[IC_H]: e_H(\hat{\theta}_H + e_H) + \alpha_H - \frac{1}{2}e_H^2 \geq e_L(\hat{\theta}_H + e_L) + \alpha_L - \frac{1}{2}e_L^2$$

The objective and zero profit conditions are identical to the observable type case, and the incentive constraints are identical to the monopsony case. It is straightforward to show that in the absence of overconfidence, neither incentive constraint is binding. Both workers are made residual claimant of their effort by setting $e_i = A$ and there is no fixed fee, $\alpha_i = 0$. Thus, competitive markets are fully efficient, even when the worker's type is his private information. When workers are overconfident, the menu of contracts firms offer when the worker's type is observable may or may not be incentive compatible. The following conditions will be useful for describing the optimal contracts and distinguishing cases:

Condition 1: $\Delta\gamma^2 \geq 2\gamma_L \Delta\theta$

Condition 2: $\Delta\gamma^2 \geq -2\gamma_H \Delta\theta$

When Condition 1 (resp. 2) is satisfied, the observable type contracts $(e_i^{CO}, \alpha_i^{CO})$ satisfy IC_H (resp. IC_L). Notice that Condition 2 is always satisfied when workers are overconfident ($\gamma_H > 0$), so it can be ignored for now, but Condition 1 may be violated when $\gamma_L > 0$.[40] The solution to P^{RS} is characterized in Proposition 2.

Proposition 2 (Perfect Competition): *When Condition 1 is satisfied, the solution to P^{RS} is:*

$$e_i^{RS} = e_i^{CO} = A + \gamma_i$$

$$\alpha_i^{RS} = \alpha_i^{CO} = -\gamma_i(A + \theta_i + \gamma_i)$$

When Condition 1 is not satisfied, the solution to P^{RS} is:

$$e_H^{RS} = A + \gamma_H = e_H^{CO}$$

$$e_L^{RS} = A + \gamma_H + \Delta\theta - X \in (A, e_L^{CO})$$

$$\alpha_H^{RS} = \alpha_H^{CO} = -\gamma_H(A + \theta_H + \gamma_H)$$

$$\alpha_L^{RS} = -(\gamma_H + \Delta\theta - X)(A + \gamma_H + \theta_H - X)$$

where $X \equiv \sqrt{2\gamma_H \Delta\theta + \Delta\theta^2}$. The effort levels of both high and low types are strictly above the efficient level.

When Condition 1 holds, the menu $(e_i^{CO}, \alpha_i^{CO})$ is incentive compatible, so it is also the RS menu. When Condition is violated, the menu $\left(e_i^{CO}, \alpha_i^{CO}\right)$ violates IC_H; high types will have an

[40] Condition 2 is used in Section 4, where I consider general differences in beliefs between firms and workers, which may include diffidence (under-confidence).

incentive to mimic low types to save on the fixed fee. Notice, however, that $(e_H^{CO}, \alpha_H^{CO})$ already maximizes the left-hand side of IC_H, so what firms must do to both attract high types and achieve incentive compatibility is to distort the contract meant for low types. The assumption that $\Delta\hat{\theta} > 0$ plays a crucial role here: though low types are overconfident about themselves, they still believe they are less productive than high types, and so the least-cost way for firms to separate low and high types is to distort the effort of low types *downward*. While this distortion does help to offset the upward distortion due to overconfidence (enablement of workers), it never completely offsets it. Hence, $e_L^{RS} \in (A, e_L^{CO})$. This has similar welfare implications to the analogous result from the monopsony case: distortions due to screening and overconfidence can offset each other in ways that improve market efficiency.

2.4.3. Competitive Equilibrium

It remains to check that the contracts derived in Proposition 2 are indeed the unique competitive equilibrium contracts. It is straightforward that when Condition 1 is satisfied, the observable type contracts are the unique competitive equilibrium contracts as no other menu yields either type of worker as much utility without forcing a firm below zero profit, i.e. there is no profitable deviation for either type of worker or any firm. When Condition 1 is violated, the competitive equilibrium may be in mixed strategies. To rule this out, consider the following definition from BT, adapted for the current setting.

Definition 1: *An incentive-compatible allocation* $\{(\tilde{e}_i, \tilde{\alpha}_i)\}_{i\in\{L,H\}}$ *is interim efficient if there exists no other incentive-compatible* $\{(e_i, \alpha_i)\}_{i\in\{L,H\}}$ *that*

(i) Pareto-dominates it: $U(e_H, \alpha_H; \hat{\theta}_H) \geq U(\tilde{e}_H, \tilde{\alpha}_H; \hat{\theta}_H), U(e_L, \alpha_L; \hat{\theta}_L) \geq U(\tilde{e}_L, \tilde{\alpha}_L; \hat{\theta}_L)$, *with at least one strict inequality.*

(ii) allows firms to break even on average:

$$q[(A-\tilde{e}_L)(\theta_L+\tilde{e}_L)-\tilde{\alpha}_L]+(1-q)[(A-\tilde{e}_H)(\theta_H+\tilde{e}_H)-\tilde{\alpha}_H]\geq 0.$$

BT prove that the RS contracts are the unique competitive equilibrium contracts *if and only if* they are interim efficient, a result which extends to this setting.[41] Now, it is a matter of deriving conditions under which the RS contracts are interim efficient. Those conditions can be found using Lemma 1.

Lemma 1: *The RS contracts are interim efficient if and only if they are the solution to the following problem:*

$$P^D: \max_{e_i,\alpha_i} q[(A-e_L)(\theta_L+e_L)-\alpha_L]+(1-q)[(A-e_H)(\theta_H+e_H)-\alpha_H]$$

$$[IR_L]: e_L(\hat{\theta}_L+e_L)+\alpha_L-\frac{1}{2}e_L^2\geq U_L^{RS}$$

$$[IR_H]: e_H(\hat{\theta}_H+e_H)+\alpha_H-\frac{1}{2}e_H^2\geq U_H^{RS}$$

$$[IC_L]: e_L(\hat{\theta}_L+e_L)+\alpha_L-\frac{1}{2}e_L^2\geq e_H(\hat{\theta}_L+e_H)+\alpha_H-\frac{1}{2}e_H^2$$

$$[IC_H]: e_H(\hat{\theta}_H+e_H)+\alpha_H-\frac{1}{2}e_H^2\geq e_L(\hat{\theta}_H+e_L)+\alpha_L-\frac{1}{2}e_L^2$$

$$[ZP]: q[(A-e_L)(\theta_L+e_L)-\alpha_L]+(1-q)[(A-e_H)(\theta_H+e_H)-\alpha_H]\geq 0$$

where

$$U_i^{RS}\equiv U(e_i^{RS},\alpha_i^{RS};\hat{\theta}_i).$$

Problem P^D can be interpreted as the problem a firm would solve when trying to determine if it had a profitable deviation from a competitive equilibrium where every other firm was offering the RS contracts. For there to be no such profitable deviation, it must be that $[IR_L], [IR_H]$ and $[ZP]$ all bind (otherwise the firm could profitably reduce either α_L or α_H), but those three

[41] See Appendix B of BT for the proof.

constraints and the binding $[IC_H]$ define the RS contracts themselves. Therefore, conditions under which the RS contracts constitute a competitive equilibrium are equivalent to conditions under which those four constraints are equalities at the solution to problem P^D, which are given as part (ii) of the following proposition.

Proposition 3 (Competitive Equilibrium): *The contracts described in Proposition 2 are the unique competitive equilibrium contracts if either:*

(i) Condition 1 is satisfied.

(ii) Condition 1 is violated and $q \leq q_H \equiv \frac{\Delta\theta + \Delta\gamma}{X} \in (0,1)$.

The intuition for the restriction that $q \leq q_H$ in part (ii) is as follows.[42] When Condition 1 fails, the interim efficient contracts may involve cross-subsidization, where firms lose money on high types in order to make more back on low types, and the competitive equilibrium will involve mixed strategies.[43] When there are too few low types in the market, firms cannot recover the losses incurred on high types without offering low types a contract that makes them worse off than the RS contract. However, any firm that does this cannot attract low types at all, meaning they are employing only high types at a loss, which is certainly not part of an equilibrium. Therefore, a market with insufficiently many low types precludes equilibrium contracts that cross-subsidize and ensures an equilibrium in pure strategies. For the purpose of doing welfare analysis, I will assume that $q \leq q_H$ hereafter.

[42] There is an analogous restriction that guarantees a pure strategy competitive equilibrium in BT.
[43] See Luz (2017) for a full characterization of the competitive equilibrium in the RS model.

2.5 Market Efficiency and Worker Welfare

2.5.1 Efficiency

Recall the definition of social surplus from a worker of type θ_i from Section 2:

$$S_i(e) \equiv A(\theta_i + e) - \frac{1}{2}e_i^2$$

and that, in the absence of overconfidence, the competitive market is efficient (since Condition 1 is always satisfied when $\gamma_i = 0$) while the monopsony market is not. This efficiency ranking corresponds standard intuition about the relationship between the degree of competition in a market and its efficiency, that more competition is better. It will turn out that this intuition does not necessarily hold in markets with both adverse selection and overconfidence.

To make efficiency comparisons in the presence of overconfidence, first define:

$$S^\zeta \equiv qS_L\left(e_L^\zeta\right) + (1-q)S_H(e_H^\zeta)$$

where $\zeta \in \{M, CO, RS\}$ indicates the relevant equilibrium contracts, M is monopsony, and CO (resp. RS) is competition where the observable type contracts are (resp. are *not*) incentive compatible. The efficiency loss in market ζ is given by:

$$L^\zeta = S^* - S^\zeta$$

where $S^* \equiv qS_L(e^*) + (1-q)S_H(e^*)$.

Proposition 4 (Market Efficiency): *For all $\gamma_i > 0$ satisfying $\Delta\hat{\theta} > 0$, the following are true:*

(i) $L^{CO} > L^{RS}$.

(ii) $L^{CO} > L^M$ if $q > q_L$.

(iii) $L^{RS} > L^M$ if $q \in (q_L, q_H)$, and the interval (q_L, q_H) is non-empty whenever $\gamma_H > 0$.

Before discussing Proposition 4, recall that the effort of high type workers is the same in all market structures: $e_H^\zeta = A + \gamma_H$ for all $\zeta \in \{M, CO, RS\}$, and the fixed payments α^ζ are purely distributive, so the efficiency properties of each market are determined entirely by e_L^ζ.

Part (i) is a comparative static result which shows that a *more severe* asymmetric information problem, in the sense of a $\Delta\theta$ large enough to violate Condition 1, can induce changes in equilibrium contract design that yield welfare gains, even in perfectly competitive markets. For small values of $\Delta\theta$, the CO contracts are incentive-compatible, and firms enable workers to the maximum extent possible. For larger enough values of $\Delta\theta$, screening types is required and can only be done by distorting the low-type effort downward, i.e., $e_L^{RS} \in (A, e_L^{CO})$, which partially offsets the upward distortion due overconfidence, improving efficiency.

Part (ii) shows that a monopsonistic market with adverse selection may be more efficient than a perfectly competitive market with no (binding) informational constraints. In either case, worker overconfidence pushes the equilibrium effort levels above the efficient levels, but the monopsonist, in an effort to extract information rent, distorts the effort of low types downward: $e_L^M \in (A, e_L^{CO})$. This has a knock-on effect of increased efficiency provided the screening distortion is not too large, which is the case when there are relatively few high types in the market from whom the monopsonist is extracting rent ($q \geq q_L$).

Part (iii) provides sufficient conditions under which the monopsonistic market is more efficient than *any* competitive market with overconfidence. The intuition for this result has to do with the different motives of competitive firms and a monopsonist. The goal of a competitive firm is to make contracts as attractive to workers as possible, since firms earn zero profit in any equilibrium. Thus, the equilibrium contracts are designed to *enable* the worker's overconfidence to the maximum extent that incentive-compatibility allows. The monopsonist, on the other hand, seeks the optimal balance between exploitation of low types, rent extraction from high types, and preserving social surplus, a portion of which ultimately constitutes her equilibrium profits. When $q \in (q_L, q_H)$, the monopsonist finds that balance when setting $e_L^M \in (A, e_L^{RS})$ which, since social surplus is strictly decreasing in e_L at all $e_L > A$, represents an unambiguous welfare improvement over the competitive market.

The results of Proposition 4 are analogous to that of Proposition 3 in BT: over-incentivization of workers can be more extensive and cause larger efficiency losses in competitive markets than monopsonistic markets. The primary difference is the mechanism. In BT, high-powered incentives induce inefficiently high effort on a measured task, which crowds out effort on an unmeasured task tasks, and increased competition accentuates this problem. In this paper, high-powered incentives come about due to overconfidence, regardless of market structure, and distortions needed to screen types can help correct the associated inefficiencies to a greater extent in monopsony than in perfect competition. Since adding overconfidence to the multitasking model of BT (or vice versa), my results are both can be viewed as a complement or alternative to theirs.

BT take the analysis a step further by developing a "full spectrum" model where the degree of competition can be varied continuously between perfect competition and monopsony. The main result from this model (their Proposition 5) is that social welfare is mound-shaped in the degree of competition and maximized at a strictly intermediate level of competition. Technically speaking, this is because which incentive constraint binds depends on the degree of competition; when competition is strong (resp. weak) enough, it is the less (resp. more) productive workers that have an incentive to misrepresent themselves. This constitutes a major difference between our models. In mine, only high types ever have an incentive to claim to be low types, and so a similar full spectrum model with overconfidence would yield a social welfare function that is monotonic in the degree of competition. Thus, I do not consider it here.

2.5.2 Worker Welfare

The welfare results of the previous subsection suggest that conventional pro-competitive policy may be counterproductive. But regulators are often willing to trade off between market efficiency and the welfare of individuals, e.g., the minimum wage, so it is worth considering whether such a trade-off exists in the presence of overconfident workers.[44] To do this, I will evaluate the preferences workers have over equilibrium contracts both at the *ex ante* stage, when workers make predictions of their equilibrium payoff while holding overconfident beliefs about their productivity, and at the *ex post* stage, where their realized payoff reflects their true productivity.[45]

Define the following:

$$\widehat{U}^{\zeta} \equiv qU\left(e_L^{\zeta}, \alpha_L^{\zeta}; \hat{\theta}_L\right) + (1-q)U(e_H^{\zeta}, \alpha_H^{\zeta}; \hat{\theta}_H)$$

$$U^{\zeta} \equiv qU\left(e_L^{\zeta}, \alpha_L^{\zeta}; \theta_L\right) + (1-q)U\left(e_H^{\zeta}, \alpha_H^{\zeta}; \theta_H\right)$$

$\widehat{U}^{\zeta}$ (resp. U^{ζ}) measures the average workers' *ex ante* (resp. *ex post*) utility from equilibrium contracts $\left(e_i^{\zeta}, \alpha_i^{\zeta}\right)_{i \in \{L,H\}}$ for $\zeta \in \{M, CO, RS\}$.

Proposition 5 (Worker Welfare): *For any* $\gamma_L, \gamma_H > 0$*:*

(i) $\widehat{U}^{CO} > \widehat{U}^{RS} > \widehat{U}^{M}$.

[44] For the sake of this exercise, I assume the regulator has all the same information as the firm. It may be unrealistic to assume a regulator has good information about a worker's true productivity within the firm, or the worker's biased belief about himself. I discuss this briefly in Section 5.

[45] Based on the analysis of Section 2, it is straightforward that a firm prefers to be a monopsonist than to compete against other firms, regardless of the severity of informational frictions or overconfidence. Competitive firms earn zero profit in any equilibrium while the monopsonist can never do worse than that.

(ii) $U^{RS} > U^{CO} > U^{M}$.

It may come as no surprise that overconfident workers prefer competitive markets (part (i)), but it is not obvious that that preference still holds *ex post*, when their mistakes are revealed to them. While workers enjoy the benefit of maximal bargaining power in the competitive markets, the forecasting errors they make in those markets also tend to be the largest.[46] It turns out that the former effect always outweighs the latter; workers are always better off in a competitive market, although they do benefit from their private information when it is relevant ($U^{RS} > U^{CO}$). While market power can result in efficiency gains, those gains are never enough to compensate workers for their loss in bargaining power. Not only is there no free lunch for a regulator, where all market participants benefit from a reduction in competition, but overconfident workers would balk at such a proposal.

2.6 Extensions

2.6.1 General Differences in Beliefs

Consider the generalization where γ_i can take on any finite real value such that $\hat{\theta}_i = \theta_i + \gamma_i > 0$.[47] This includes the possibility that either $\gamma_i < 0$, which represent diffidence (under-confidence), as well as $\Delta\hat{\theta} < 0$, where the differences in beliefs cause firms and workers to rank types differently (since firms know that $\Delta\theta > 0$).[48] This generalization involves many cases, not all of which are particularly rich in content beyond that of the baseline model. Instead of discussing all possible cases here, I will outline the basic logic of the equilibrium contracts in both monopsony

[46] Forecasting errors are given by $E^{\zeta} \equiv e_L^{\zeta}\gamma_L$. Therefore, $E^{M} < \min\{E^{CO}, E^{RS}\} = E^{RS}$ is equivalent to $e_L^{M} < e_L^{RS}$, which is a necessary condition for $L^{M} < L^{RS}$, the case of main interest in this paper.

[47] If either $\hat{\theta}_i < 0$, that type cannot be motivated to exert effort.

[48] I omit the knife-edge case of $\Delta\hat{\theta} = 0$, wherein types cannot be separated.

and perfect competition and provide some simple sufficient conditions for results similar to Proposition 4 to hold. Before moving on to the discussion of the equilibrium contracts, note that the definition of social surplus ($S_i(e_i)$), the efficient level of effort ($e_i = A$), observable type level of effort ($e_i^{CO} = A + \gamma_i$), and maximized social surplus (S_i^*) are all unaffected by this more general setting.

In monopsony, the way workers rank types, i.e., the sign of $\Delta\hat{\theta}$, determines which type of worker will have an incentive to misrepresent himself to the firm which, in turn, determines which type's effort the monopsonist distorts downward in order to extract information rent. Importantly, this downward distortion only improves efficiency if it is offsetting an upward distortion due to overconfidence. Therefore, there are two cases where monopsony power improves efficiency. The first is from Section 2, where $\Delta\hat{\theta} > 0$ and $\gamma_L > 0$. When $\Delta\hat{\theta} > 0$ high-type's incentive constraint is binding, and the low-type's effort is distorted downward to relax it. The second is where $\Delta\hat{\theta} < 0$ and $\gamma_H > 0$. In this case, the low-type's incentive constraint is binding since workers believe that *low types* are the more productive type. The monopsonist distorts e_H^M downward to extract rent, which offsets the upward distortion due to $\gamma_H > 0$. Indeed, high type effort in this scenario is given by:

$$e_H^M = A + \gamma_H + \frac{q}{1-q}\Delta\hat{\theta} < A + \gamma_H = e_H^{CO}.$$

In perfect competition, what determines which type of worker may have an incentive to misrepresent his productivity depends on Conditions 1 and 2 given in Section 2:

Condition 1: $\Delta\gamma^2 \geq 2\gamma_L\Delta\theta$

Condition 2: $\Delta\gamma^2 \geq -2\gamma_H\Delta\theta$

Recall that when Condition 1 (resp. 2) is satisfied, the observable type contracts (e_i^{CO}, α_i^{CO}) satisfy IC_H (resp. IC_L). It is also clear that $\gamma_L > 0$ (resp. $\gamma_H < 0$) is a necessary condition these contracts to violate Condition 1 (resp. Condition 2), since $\Delta\theta > 0$. Using the same logic as in Section 2, the binding incentive constraint will determine which type must be offered the observable type contract, but in which direction the other type's effort gets distorted to achieve incentive compatibility depends on the sign of $\Delta\hat{\theta}$. Finally, whether that distortion improves efficiency or not depends on in what direction the distortion due to disagreement goes.

Laying these various conditions on top of each other yields several different cases, but ultimately most of them can be disregarded, typically because there are no conditions under which the RS contracts are interim efficient.[49] Analysis of the case that remain yields Proposition 6.

Proposition 6 (Market Efficiency with General Disagreement): *Each of the following is a sufficient condition for* $L^M < \min\{L^{RS}, L^{CO}\}$:

(i) Condition 1 and 2 are both satisfied, $\Delta\hat{\theta} > 0$ *and* $\gamma_L > 0$.

(ii) Condition 1 and 2 are both satisfied, $\Delta\hat{\theta} < 0$ *and* $\gamma_H > 0$.

(iii) Condition 1 is violated,$\gamma_i > 0$, $\Delta\hat{\theta} > 0$, *and* $q \in (q_L, q_H)$.

(iv) Condition 2 is violated, $\gamma_i < 0$, $\Delta\hat{\theta} < 0$, *and* $q < q'_H$ *where* $q'_H \equiv \frac{\sqrt{\Delta\theta^2 - 2\gamma_L\Delta\theta} + \Delta\hat{\theta}}{\sqrt{\Delta\theta^2 - 2\gamma_L\Delta\theta}} \in (0,1)$.

The underlying intuition for each of parts (i)-(iii) is the same as in the baseline model: distortions in equilibrium effort levels due to the worker's overconfidence may be mitigated most effectively by screening distortions in monopsonistic markets. Indeed, part (iii) includes the case of the baseline model. The intuition from part (iv) is a bit different. In this case, $\gamma_H < 0$ (by implication of Condition 2 being violated), so we have $e_H^{CO} < e_H^*$, and since $\Delta\hat{\theta} < 0$, it is $[IC_L]$ that binds in both market structures. Thus, screening distortions will have high types taking less

[49] See Appendix B for details.

effort than in the observable type case, i.e., $e_H^{RS}, e_H^{M} < e_H^{CO} < e_H^{*}$, so these distortions actually exacerbate those due to the high-type worker's diffidence. For the monopsonistic market to be more efficient than the perfectly competitive market, this should happen to a lesser extent under monopsony, which is true when $e_H^{RS} < e_H^{M}$, and since $e_H^{RS} < e_H^{CO}$, there exists a q small enough, call it q_H', such that this is the case. Finally, while the qualitative nature of the distortions is different in (iv), the underlying intuition that the monopolist has an indirect incentive to prevent drastic inefficiencies that a competitive firm lacks still holds.

2.6.2 Non-linear Contracts

In the baseline model, I use linear contracts, *á la* BT, for their simplicity and ease of interpretation. One might be concerned that the results are idiosyncratic to this case, but allowing for fully nonlinear contracts in the baseline model creates a technical issue: it is possible for the monopsonist to turn the worker into a "Dutch Book" and earn an arbitrarily large profit. This is because, for a given level of effort e, the worker believes his output will be $(\hat{\theta}_i + e)$, but his realized output will then be $(\theta_i + e) < (\hat{\theta}_i + e)$. The monopsonist could offer the worker an arbitrarily large payment when $(\hat{\theta}_i + e)$ realizes and charge him an arbitrarily large fee when any other output realizes (which it always will), satisfying all constraints. Since this outcome is sufficiently unrealistic as to rule out, considering nonlinear contracts requires altering the model.

Suppose that the worker's type, θ, now represents his marginal cost, i.e., his disutility of effort is given by $c(e;\theta) = \frac{\theta}{2}e^2$, and his output is given by $Y = Ae$.[50] Firms compensate workers with a transfer $t(e)$, so profit is given by:

[50] In this version of the model, firms can observe the worker's effort, but not his cost, which means the worker may still have an incentive to over/under-state his cost to earn an information rent.

$$\Pi(t(e); e, \theta) = Ae - t(e)$$

and the worker's payoff is given by:

$$U(e; t(e), \theta) = t(e) - \frac{\theta}{2}e^2.$$

Firms believe that $\theta \in \{\theta_L, \theta_H\}$ where $\theta_H > \theta_L > 0$, while workers believe that $\theta \in \{\hat{\theta}_L, \hat{\theta}_H\}$ where $\hat{\theta}_H > \hat{\theta}_L > 0$, $\hat{\theta}_i \equiv \theta_i - \gamma_i$ and $\theta_i > \gamma_i \geq 0$. While firms and workers agree to disagree on the true values θ can take, they agree on the distribution of types, where q is the probability that the worker is a *high type* ($\theta = \theta_H, \hat{\theta}_H$) and $(1 - q)$ is the probability that the worker is a *low type* ($\theta = \theta_L, \hat{\theta}_L$).[51] Lastly, social welfare and the efficient effort level are each defined analogously in this new setting:

$$S_i(e_i) \equiv Ae_i - \frac{\theta_i}{2}e_i^2 \Rightarrow e_i^* = \frac{A}{\theta_i}$$

$$S_i^* \equiv S_i(e_i^*) = \frac{A^2}{2\theta_i}.$$

A great deal of the exposition in Section 2 applies to this new model, so I will describe the important features of the optimal contracts here, state and discuss the main welfare results, and direct the interested reader to Appendix B for the formal analysis. For now, it is helpful to point out that in either monopsony or perfect competition, the effort level firms will offer workers when their types are observable are given by:

$$e_i^{CO} = \frac{A}{\hat{\theta}_i} > \frac{A}{\theta_i} = e_i^*$$

In the monopsony case, carrying the assumption that $\Delta\hat{\theta} > 0$ over from the baseline model, the only thing non-standard about the monopsonist's problem is the worker's overconfidence. Efficient (low type) workers take their observable type effort, $e_L^M = e_L^{CO}$ and earn an information

[51] Thus, as in the baseline model, q is the proportion of less "able" workers.

rent given by $\frac{\Delta\hat{\theta}}{2}(e_H^M)^2$. Inefficient (high type) workers' effort is distorted below the observable type level, $e_H^M < e_H^{CO}$, to limit the high type's information rent, and earn no rent themselves.

In competitive markets, the analysis turns out to be simpler than in the baseline model. Neither incentive constraint binds when firms offer the observable type contracts, and since the associated effort levels maximize the utility of each type of worker among those that yield the firm non-negative profit, they also constitute the unique competitive equilibrium effort levels. Since the effort of low types is the same in both market structures, efficiency comparisons will simplify down to comparison between e_H^M and e_H^{CO}. Using the definition of the efficiency loss L^ζ given in Section 3, Proposition 7 restates the main results from Proposition 4 in this new context.

Proposition 7 (Market Efficiency with Nonlinear Contracts): *For all* $\gamma_i > 0$ *satisfying* $\hat{\theta}_H > \hat{\theta}_L > 0$, $L^{CO} > L^M$ *if* $q > \tilde{q}_L \equiv \frac{\Delta\hat{\theta}}{\Delta\hat{\theta}+\gamma_H}$.

All the same intuition from the baseline model remains valid. Competitive firms enable workers' overconfidence, offering them contracts with effort levels above the efficient level. The monopsonist requests the same effort level from low types as offered by competitive firms, but a strictly lower effort level from high types. This reduction in the high type effort is meant to extract rent from low types, something only a profitable firm has any incentive to do. Thus, the monopsonist's profit-maximizing behavior helps offset the exploitative upward distortions in high-type effort, which improves efficiency as long as distortions due to rent extraction are not too large ($q > \tilde{q}_L$).

2.7 Conclusion

This paper studies how asymmetric information, overconfidence, and the degree of market competition interact to influence market efficiency. It has shown that, in both competitive and monopsonistic markets, distortions due to overconfidence and adverse selection can offset each other, in which case markets with asymmetric information problems turn out to be more efficient than those without them. It has also shown that these offsetting distortions appear for different reasons in competitive and monopsonistic markets, that these differing motives lead to differently sized distortions, and that the monopsonistic market may at times be more efficient than the competitive market, provided the labor market is sufficiently diverse in terms of worker productivity. These efficiency results are qualitatively similar to those of BT but derived under different assumptions, thus providing both an alternative and complementary explanation to theirs.

One direction in which this analysis could be extended is in adopting alternative definitions of overconfidence. There are many definitions used in the literature. Some examples are overly precise beliefs (Grubb 2009), overestimation of the persistence in one's own preferences (Eliaz and Spiegler 2006), overestimation in the quality of one's information (Goel and Thakor 2008), or overestimation of success in some venture (De La Rosa 2011). Ideally, the results ought not depend on the definition of overconfidence used in the model. Unfortunately, adapting the model to other forms of overconfidence can require it be completely rewritten, or otherwise introduces substantial technical difficulties, so it is left for future work.

The model also lends itself to policy analysis, as in BT. They consider bonus caps, total earnings caps, and taxation as potential policy tools for correcting the inefficiencies that result from high-powered incentive contracts. They find that bonus caps can restore perfectly competitive markets to full efficiency but argue that it may not be possible for the regulator to distinguish between bonus pay and base pay. When only total earnings can be observed, BT suppose that firms may substitute towards non-monetary payments that the different types of workers enjoy differentially. They show that when the firms have access to non-monetary rewards that are a close enough substitute to monetary payments, a binding earnings cap can improve welfare. Introducing regulation into a model that involves disagreement can be substantially more complicated than introducing it into a “standard” model. As I argued earlier, there is reason to believe that firms may have the best information about workers’ true productivities within the firm, in which case the regulator would need to illicit that information from firms in order to construct optimal policies. Even in a “reduce form” model where the regulator’s beliefs are taken as given, it is not obvious what beliefs the regulator holds, and allowing for any belief immediately results in a large number of cases. A thorough analysis of these issues and their implications are beyond the scope of this paper and left for future research.

Chapter 3

Overconfidence in Contract Theory

Abstract: I catalogue several definitions of overconfidence found in the behavioral contract theory literature and apply them each to a standard information gathering contracting model. The effect overconfidence has on the agent's willingness to sign the contract, to gather information, and on the principal's profit varies qualitatively across definitions. The results demonstrate that overconfidence can have countervailing effects on multifaceted incentive problems. Some implications for future theoretical and empirical research are discussed.

3.1 Introduction

By incorporating overconfidence into contract-theoretic models, researchers have been able to explain a wide range of real-world phenomena. These phenomena include, for example, wage compression and gender pay discrepancies (Fang and Moscarini (2005), Sautmann (2011), Santos-Pinto (2012)), the structure of contracts for subscription-based services like cell phone coverage (Grubb (2009)), the promotion of overconfident workers within organizations (Goel and Thakor (2008)), and the inverse relationship between risk and coverage in insurance markets (Sandroni and Squintani (2013)). This literature has also improved our understanding of the effectiveness of policy in markets with overconfident consumers (see, for example, Sandroni and Squintani (2007) or Manove and Padilla (1999)).

While the literature has produced many valuable insights, the lack of a single standard model or mathematical definition of overconfidence makes it difficult to put those insights into appropriate context, to compare models of overconfidence to one another, and to summarize the literature's findings. In fact, in some cases the definition of overconfidence used in one model is entirely inapplicable another, making their predictions incomparable. So far, all that has been done to address this issue is to distinguish between *overoptimism* and *overprecision* (See Grubb (2015), Herz et al (2014), or Malmedier and Taylor (2015)), but there are still many more definitions of overconfidence than these two terms can precisely describe on their own.

The goal of this paper is to catalogue various definitions of overconfidence found in the behavioral contract theory literature and to apply them to a single baseline model where their effects can be assessed and compared. To do this, I adapt the model in Downs (2020a), which is itself a simplified version of the standard information gathering contracting model of Crémer, Khalil, and Rochet (1998). In these models, a principal (she) hires an agent (he) to exert costly effort to gather information about the cost of implementing a project (or producing a good) before he engages in production itself. Since the agent's effort, whether or not he learns and what he learns are all his private information, the principal must provide him with incentives to bother gather information and to reveal it to her. This setting is an ideal baseline for studying different forms of overconfidence since it involves adverse selection, moral hazard, and the acquisition and processing of information, each of which may be influenced by behavioral biases.

The paper is organized as follows. In Section 2, I state the baseline model, which is a version of the model in Downs (2020a) that has been generalized to accommodate several different definitions of overconfidence from the literature. Section 3 lists definitions of overconfidence found in the literature and applies them to the baseline model. For each case, there are three questions that I will answer: (i) how does the agent's overconfidence affect his willingness to sign the contract, (ii) how does the agent's overconfidence affect his willingness to gather information, and (iii) how does the principal's payoff vary with the agent's overconfidence?[52]

The answers to these questions provide some important lessons about the effects of overconfidence in contracting models. First, in *almost* every definition of overconfidence considered below, overconfidence makes the agent more willing to sign the contract, since it causes him to overestimate his payoff from doing so. This result is consistent with simple intuition as well as most of the results found in the literature. Second, whether overconfidence makes the agent more or less willing to gather information depends on whether his overconfidence affects his subjective expected payoff from *not* gathering information. Types of overconfidence that make

[52] To abstract from some technical minutia and focus on intuition, I will say that an agent is more (resp. less) willing to gather information when either (i) a binding information gathering incentive constraint is relaxed (resp. tightened) by overconfidence or (ii) all information gathering incentive constraints are non-binding and can be satisfied for a larger (resp. smaller) set of parameter values with an overconfident agent. Additionally, I will restrict attention to cases where there exists at least one contract that induces information gathering and omitting from the discussion contracts that do not induce information gathering.

remaining uninformed more (resp. less) attractive to the agent will harm (resp. improve) information gathering incentives. While it is obvious that an agent will have a greater incentive to deviate when his deviation becomes more attractive, this conclusion can help provide researchers who are studying overconfidence with a better sense of how it should be defined in order for their modeling choices to accurately reflect the realities of the situations they study. Thus, this paper makes a methodological contribution to the theory on contracting with overconfidence agents, as well serving as a review of definitions of overconfidence used in that literature.

Finally, in Section 4 I discuss how overconfidence relates to time-inconsistency, another behavioral bias that is commonly applied to contracting models and provide some concluding remarks.

3.2 Baseline Model

A principal (she) hires an agent (he) to implement a project of size $q \in \{0,1\}$. The principal values the project according to $V(q) = Vq$. The agent's cost of implementing the project is βq where $\beta \in \{\beta_L, \beta_H\}$, $\beta_H > \beta_L > 0$ and $\Delta\beta \equiv \beta_H - \beta_L$. Project returns are such that it is *ex post* efficient to implement the project if and only if costs are low: $\beta_H > V > \beta_L > 0$.

Prior to contracting, the principal and agent are each endowed with beliefs about the distribution of the implementation cost β. Let their common prior be that $\Pr(\beta = \beta_L) = \eta \in (0,1)$ and $\Pr(\beta = \beta_H) = 1 - \eta$. At the time of contracting, neither the principal nor the agent knows the value of β. After the contract is signed, the agent may exert effort $a \in \{0,1\}$ at utility cost ca, to obtain a signal σ that is correlated with the implementation cost β. Specifically, the agent obtains the signal with probability $p + \pi a$ where $p \in [0,1)$ and $\pi \in (0,1 - p]$ and obtains no signal with probability $1 - p - \pi a$. The signal takes values $\sigma \in \{\sigma_L. \sigma_H\}$ where

$\Pr(\beta = \beta_i | \sigma = \sigma_i) = r$, where $r > \frac{1}{2}$ without loss of generality. The agent's effort a, whether he obtains the signal σ or not, and what value of σ he observes, are all the agent's private information. The principal compensates the agent for the costs of implementation and information gathering effort with a monetary transfer t. Their payoffs are:

$$\Pi(q, t) = Vq - t$$

$$U(q, t; \beta) = t - \beta q - ca$$

Each player has a reservation payoff that is normalized to zero. A contract is a menu $(t_i, q_i)_{i \in \{L,H,U\}}$, where $i = L$ indicates an agent who has observed σ_L, $i = H$ an agent who has observed σ_H, and $i = U$ an agent who has not observed the signal.[53] The timing of the game is as follows:

- $t = 1$: The principal offers a menu of contracts $(t_i, q_i)_{i \in \{U,L,H\}}$ to the agent, who accepts or rejects. If he rejects, the game ends and the players receive their reservation payoffs. If he accepts, the game proceeds to $t = 2$.
- $t = 2$: The agent privately exerts information gathering effort a. He privately observes either the signal σ or nothing.
- $t = 3$: The agent implements a project of size q, which is verifiable and delivered directly to the principal.
- $t = 4$: Contractual payments are made, and payoffs are realized.

The optimal contract involves two basic cases, one where it induces the agent to gather information and one where it does not, but each of these cases involves subcases of their own.

[53] Since the implementation decision is binary, it will be the case that uninformed agents are pooled with either low-cost or high-cost agents in equilibrium. Which type of informed agents that uninformed agents are pooled with has no qualitative effects on the predictions of the models provided in the following section.

Since describing all possible outcomes across the entire parameter space is neither the goal of this paper nor will it contain any new insights, I will make the following assumptions:[54]

(A1): It is efficient to gather information under the principal's beliefs and complete information.

(A2): The parameters are such that the optimal contract induces information gathering and yields the principal her complete information payoff under agreement (when the agent is not overconfident).

Under these assumptions, the optimal contract offered to an agent who is not overconfident will induce him to gather information, to implement the project when he observes σ_L and to not implement it when he observes σ_H, all at the minimum possible cost to the principal.[55] Thus, for any definition of overconfidence used in the following section, there is an open set in the parameter space wherein inducing information gathering by an overconfident agent is also optimal. These assumptions allow me to restrict attention to contracts that induce information gathering, making it possible to assess the effect of overconfidence on the agent's willingness to do so. They also fix the principal's profit in the absence of overconfidence at a natural benchmark, her first-best profit, so that any benefit or cost of the agent's overconfidence can be easily identified.[56]

3.3 Definitions of Overconfidence

The structure of this section is as follows: each subsection starts with a description of one type of overconfidence found in the literature, as well as in what papers it is used. Then, since

[54] See, for example, Iossa and Martimort (2015) for a model where information gathering may fail and Gervais, Heaton, and Odean (2011) for one where the information the agent gathers is an imperfect signal of the true state.

[55] There are only two cases in the following section where the agent could be uninformed in equilibrium. In each of those cases, the qualitative effects of overconfidence do not depend on whether an uninformed agent implements the project ($q_U = 1$) or not ($q_U = 0$).

[56] While it is possible to state mathematical versions of these assumptions for the baseline model, they are more useful when put into the context of each specialized model described below. Mathematical version of assumptions (A1) and (A2) corresponding to each definition of overconfidence are included in the appendix.

each model is slightly different, I will make simplifying assumptions on parameters that do not meaningfully interact with the specific definition of overconfidence for that model. For instance, in some cases I will assume that $p = 0$, which means that exerting effort is necessary (but maybe not sufficient) for the agent to obtain a signal. Letting $p = 0$ and $\pi = 1$ makes information gathering deterministic in that he obtains a signal if and only if he exerts effort and letting $r = 1$ means the signal σ is perfectly correlated with the true implementation cost β. Following these parameter assumptions is the mathematical definition of the type of overconfidence under discussion when applied to the baseline model. Finally, I will state the results for each model, i.e., the effect overconfidence has on the agent's incentive to sign the contract and to gather information, as well as the effect on the principal's payoff. Each subsection concludes with a discussion of the intuition for those results. The proofs for each case can be found in the appendix, along with the associated mathematical version of assumptions (A1) and (A2).

3.3.1 The agent underestimates the cost of information gathering effort (Santos-Pinto (2012)).[57]

Let $p = 0, \pi = 1$ and $r = 1$. *Ex ante*, the agent believes that his cost of information gathering effort is $\hat{c}a$ where $\hat{c} \in (0, c]$.

Proposition 1: *When the agent is overconfident in the sense that he underestimates his information gather cost:*

(i) He is relatively more willing to sign the contract.

(ii) He is relatively more willing to exert effort gathering information.

[57] Santos-Pinto (2012) presents a Spence job market signaling model with some overconfident workers and some underconfident workers. Those that are overconfident underestimate their marginal cost of education.

(iii) *The principal's profit is strictly increasing in the degree of overconfidence, i.e., strictly decreasing in $\hat{c}$.*

This situation is straightforward: in underestimating his cost of information gathering effort, the agent overestimates his payoff from signing any contract that induces information gathering, which makes him more willing to participate. It also makes him more willing to gather information since a lower cost of effort makes shirking relatively less attractive. The improved incentives for both participation and information gathering benefit the principal in the form of increased profits.

3.3.2 The agent overestimates the probability of having a low implementation cost (Downs (2020a), Fang and Moscarini (2005), Landier and Thesmar (2009)).

Let $p = 0, \pi = 1$ and $r = 1$. Let $\eta \in (0,1)$ be the principal's prior belief that $\beta = \beta_L$ and $\hat{\eta} \in (\eta, 1)$ be the agent's prior belief that $\beta = \beta_L$.

Proposition 2: *When the agent is overconfident in the sense that he overestimates the probability of having low implementation costs:*

(i) *He is relatively more willing to sign the contract.*

(ii) *He is relatively* ***less*** *willing to gather information.*

(iii) *The principal's profit is non-monotonic in the degree of overconfidence, increasing for low levels of overconfidence, and decreasing for high levels.*

The agent's underestimation of his expected implementation cost affects his decision making in two ways. First, his overconfidence makes him more willing to accept a given contract. This is because, to motivate information gathering, the agent must be rewarded for not overstating his costs, i.e., when he reveals that he has observed σ_L. Overconfidence makes him overestimate the probability of receiving this reward, making him more willing to participate (the *participation effect*). Second, overconfidence makes the agent less willing to gather information: making an uninformed report that his implementation cost is low is only costly when that report turns out to be wrong, and an overconfident agent underestimates the probability that his (best) uninformed report is wrong (the *incentive effect*). This effect harms the principal.

Overall, either one of these effects may be larger. While the agent's level of overconfidence is low, the participation effect, which is a first order effect, is larger, and the principal's profit is increasing in the degree of overconfidence. When the agent's level of overconfidence is high enough, the incentive effect, a second order effect, overtakes the participation effect and the principal's profit is decreasing in the degree of overconfidence.[58]

3.3.3 The agent underestimates his realized implementation cost (Downs (2020b)).

Let $p = 0, \pi = 1$ and $r = 1$. The agent believes that $\beta \in \{\hat{\beta}_L, \hat{\beta}_H\}$, where $\hat{\beta}_i = \beta_i - \gamma_i$ for $\gamma_i \in [0, \beta_i), \Delta\hat{\beta} \equiv \hat{\beta}_H - \hat{\beta}_L > 0,$ and $\hat{\beta}_A = \eta\hat{\beta}_L + (1-\eta)\hat{\beta}_H$.[59] After he observes the signal σ_i, the agent believes that $\beta = \hat{\beta}_i$.

Proposition 3: *When the agent is overconfident in the sense that he underestimates his realized implementation cost:*

[58] As I show in Downs (2020a), for high enough levels of overconfidence information gathering becomes infeasible; no contract can satisfy both information gathering incentive constraints. When this is true, overconfidence benefits the principal since she will ask the agent to implement the project and compensate him for its expected cost, which he underestimates. This case is omitted from the discussion since the agent does not gather information.

[59] Letting $\Delta\hat{\beta} \equiv \hat{\beta}_H - \hat{\beta}_L > 0$ is an assumption that does not need to hold, but has some intuition behind it: the agent may be wrong about the true values of the implementation costs but he *ranks* the costs in the same way the principal does. I discuss alternatives to this assumption in Downs (2020b).

(i) *Increases in the agent's overconfidence when he is a low type make him more willing to participate, more willing to gather information, and increase the principal's profit.*

(ii) *Increases in the agent's overconfidence when he is a high type have no effect on his willingness to participate, make him* ***less*** *willing to gathering information, and have no effect on the principal's profit.*

While the agent's overconfidence only manifests itself at the implementation stage, his anticipation of drawing a realized cost from the set $\{\hat{\beta}_L, \hat{\beta}_H\}$ affects his decision-making at the contracting and information gathering stage. Consider first increases in γ_L, i.e., increases in the overconfidence of the agent when he turns out to have low costs. At the contracting stage, this causes the agent to overestimate his payoff from a given contract, since he anticipates implementing the project at a lower cost when he observes σ_L. This makes him more willing to participate. At the information gathering stage, the agent has two possible deviations: to shirk and claim to have observed σ_L to or shirk and claim to have observed σ_H. If he were to shirk and claim to have observed σ_L, he will have to implement the project and expects to have a low implementation cost $\hat{\beta}_L$ with the same probability η as he would have if he were to exert information gathering effort. Thus, overconfidence has no effect on his incentive to deviate in this way. However, the same is not true with respect to his other available deviation. Underestimating β_L increases his expected payoff from exerting information gathering effort but has no effect on his payoff from claiming to have observed σ_H, thus increasing his willingness to gather information. The improved incentive to participation and gather information makes the principal better off.

Now consider the effect of increases in γ_H, i.e., increases in the agent's overconfidence when his realized costs are high. Since the agent will not implement the project when he observes

σ_H, changes in γ_H have no effect on his participation decision. However, when he considers shirking on information gathering and claiming to have observed σ_L, he will be implementing the project at an expected cost $\hat{\beta}_A \equiv \eta\hat{\beta}_L + (1-\eta)\hat{\beta}_H$. Since his expected cost is decreasing in γ_H, increases in γ_H make this deviation becomes more attractive to the agent, making him less willing to gather information. In spite of this, the effect on the principal's profit is more subtle: since the principal and agent agree on the likelihood of either state, the distribution of transfers has no effect on the principal's profit provided they satisfy all constraints. This means the optimal contract that induces information gathering is *not* unique, and so reductions in the agent's willingness to gather information do not affect the principal's profit for any contract that induces information gathering.

3.3.4 The agent overestimates his effort's contribution to success in gathering information (Iossa and Martimort (2015), De La Rosa (2011), Gervais and Goldstein (2007), Kim (2015). Santos-Pinto (2008)).

Let $p = 0$ and $r = 1$. The agent believes that the probability of obtaining the signal σ is $\hat{\pi}a$ where $\hat{\pi} \in [\pi, 1]$.

Proposition 4: *When the agent is overconfident in the sense that he overestimates his effort's contribution to success in information gathering:*

- *(i) The agent is relatively more willing to sign the contract.*
- *(ii) The agent is relatively more willing to gather information.*
- *(iii) The principal's profit is strictly increasing in the degree of overconfidence, i.e., strictly increasing in $\hat{\pi}$.*

The intuition is as follows: to be motivated to exert effort gathering information, the agent must be rewarded when he is informed relative to when he is uninformed. Since the agent overestimates the probability of becoming informed conditional on exerting effort ($\hat{\pi} > \pi$), he overestimates the probability of obtaining that reward. This causes him to overestimate his equilibrium payoff where he exerts effort and has no effect on his payoff when shirking. This makes him both more willing to sign the contract and more willing to exert effort gathering information. Since incentives are improved on both fronts, overconfidence unambiguously benefits the principal.

3.3.5 The agent overestimates the precision of the signal (Gervais, Heaton, and Odean (2011), Goel and Thakor (2008), Herwig and Muller (2016)).

Let $p = 0$ and $\pi = 1$. The agent believes that $\Pr(\beta = \beta_i | \sigma = \sigma_i) = \hat{r} \in [r, 1]$.

Proposition 5: *When the agent is overconfident in the sense that he overestimates the precision of his signal:*

- *(i)* *He is relatively more willing to sign the contract.*
- *(ii)* *He is relatively more willing to gathering information.*
- *(iii)* *The principal's profit is strictly increasing in the degree of overconfidence, i.e., strictly increasing in* $\hat{r}$.

The main issue with a signal that is imperfectly correlated with the true state ($r \in \left(\frac{1}{2}, 1\right)$) is that sometimes the signal is wrong, and the agent either misses out on the higher payoff he earns from implementing a low-cost project or he earns a low (possibly negative) payoff from erroneously implementing a high-cost project. At the contracting stage, the principal must compensate the agent for these circumstances in expectation, which is costly to her, but an overconfidence agent *underestimates* the probability of these mistakes, making him cheaper to employ. His willingness to sign the contract is increased by his overconfidence.

Unlike previous cases, the precision of the signal influences both the agent's incentive to gather information *and* his incentive to truthfully reveal that information. It turns out that his overestimation of the accuracy of the signal makes him more willing to do both. The agent's expected cost of implementation conditional on observing σ_L is decreasing in the precision of the signal, meaning a more precise signal makes him less interested in misreporting that he has observed σ_H. Similarly, the agent's expect cost of implementation conditional on observing σ_H is *increasing* in the precision of the signal, meaning a more precise signals makes him less interested in misreporting σ_L. Overall, overestimating the precision of the signal improves truth-telling incentives.

As for information gathering incentives, the precision of the signal has no effect on the agent's payoff should he shirk, since he receives no signal at all, but increases his payoff (from a fixed contract) from gathering information by the same logic as was just described in explaining the effect on his participation incentives. Therefore, overestimating the precision of the signal also improves information gathering incentives. Since overconfidence makes the agent more willing to participate, gather, and reveal information, the agent's overconfidence unambiguously benefits the principal.

3.3.6 The agent overestimates his chance of success independent of effort (De La Rosa (2011), Bond and Newman (2009)).

Let $r = 1$. The agent believes that his probability of successfully gathering information is $\hat{p} + \pi a$ where $\hat{p} \in [p, 1]$ and $\pi < 1 - \hat{p} \leq 1 - p$.

Proposition 6: *When the agent is overconfident in the sense that he overestimates the probability of successfully gathering information independent of effort:*

(i) *He is relatively more willing to sign the contract.*

(ii) *His willingness to gather information is unaffected by his overconfidence.*

(iii) *The principal's profit is strictly increasing in the degree of overconfidence, i.e., strictly increasing in $\hat{p}$.*

With respect to participation incentives, the logic is similar to that in 3.4.: the agent is rewarded for being informed in order to incentivize his information gathering effort, and his overconfidence causes him to overestimate the probability of earning that reward, so he is relatively more willing to sign the contract. With respect to information gathering incentives, his overconfidence has no effect, since his overconfidence does not alter *the incremental* effect of his effort on successfully gathering information, which is given by π. Since information gathering incentives are not affected by overconfidence, but participation incentives are improved, overconfidence unambiguously benefits the principal.

3.3.7 The agent sometimes misinterprets σ_H as σ_L (Manove and Padilla (1999)), leading him to, with some probability, believe he has low costs when in fact his costs are high (Sandroni and Squintani (2007, 2013)).

Let $p = 0, \pi = 1$ and $r = 1$. Suppose that with probability μ the agent is an "optimist", who observes $\hat{\sigma}$ rather than σ, and that $\Pr(\hat{\sigma} = \sigma_L) = 1$, while with probability $(1 - \mu)$ he is a "realist", and observes σ, and that all agents believe they are realists.

A consequence of this is that, at the implementation stage, the probability that the agent *correctly* believes that $\beta = \beta_L$ is η, the probability that the agent *correctly* believes that $\beta = \beta_H$ is $(1 - \eta)(1 - \mu)$, and the probability that the agent *incorrectly* believes $\beta = \beta_L$ when in fact $\beta = \beta_H$ is $(1 - \eta)\mu$. That is, with probability $(1 - \eta)\mu$ the agent is overconfident in that he believes his cost is β_L when it is actually β_H.

Proposition 7: *When the agent is overconfident in the sense that he sometimes believes his cost is low when it is high:*

(i) *His willingness to sign the contract is unaffected by his overconfidence.*

(ii) *His willingness to gather information is unaffected by his overconfidence.*

(iii) *The principal's profit is strictly increasing in the probability of an agent with high costs believing he has low costs, i.e., strictly increasing in μ.*

At the contracting and information gathering stages, an optimistic agent is essentially identical to a realistic agent: while he will eventually misinterpret a signal that his costs are high as one that indicates his costs are low, he is not yet overconfident. Thus, overconfidence has no effect on his incentives at all in this model. However, this does not mean the principal is not affected by his overconfidence. Since an optimist will believe his costs are lower than they will turn out to be implementation stage, optimists will behave like realists with low implementation costs. With respect to the principal's payoff, this is equivalent to an increase in the *ex-ante* probability that she is contracting with a low-cost agent, which increases here payoff.

3.3.8 The agent has an overly *precise* belief about the distribution of his implementation cost (Grubb (2009))

Let $p = 0, \pi = 1$ and $r = 1$. The agent believes that $\beta \in \{\hat{\beta}_L, \hat{\beta}_H\}$ where $\beta_H > \hat{\beta}_H > \hat{\beta}_L > \beta_L > 0$ and $\hat{\beta}_A \equiv \eta\hat{\beta}_L + (1-\eta)\hat{\beta}_H = \eta\beta_L + (1-\eta)\beta_H \equiv \beta_P$. That is, the principal and agent share a common prior on the *expected* implementation cost, but the agent underestimates the variance. Letting $\hat{\beta}_H - \hat{\beta}_L \equiv \Delta\hat{\beta}$, then $\Delta\beta - \Delta\hat{\beta} \equiv \delta$ is a measure of the degree of the agent's overconfidence.

Proposition 8: *When the agent is overconfident in the sense that he underestimates the variance of his implementation cost:*

- *(i)* *He is relatively* ***less*** *willing to sign the contract.*
- *(ii)* *He is relatively* ***less*** *willing to gather information.*
- *(iii)* *The principal's profit is strictly* ***decreasing*** *in the degree of overconfidence, i.e., strictly decreasing in* δ.

In order to agree with the principal on the expected implementation cost, but *underestimate* the variance, it must be that the agent *overestimates* β_L. Since, in equilibrium, the agent will only implement the project when he observes σ_L, his overestimation of β_L makes him underestimate his payoff from signing the contract, so he is relatively less willing to do so. This same logic explains his reduced incentive to gather information as well since the agent's expected payoff from remaining uninformed is not affected by his overconfidence; he does not implement the project when misreporting σ_H and implements under an unbiased expected cost when misreporting σ_L. Since overconfidence harms incentives for both participation and information gathering, it must make those actions more costly to induce, meaning the principal's profit is reduced by overconfidence.

3.4 Discussion and Conclusion

3.4.1 The relationship between hyperbolic discounting and overconfidence

There are several papers that discuss agents who are time inconsistent as a result of quasi-hyperbolic discounting, and who may or may not be partially or fully sophisticated with respect to their time inconsistency (see, for example, Laibson (1997), O'Donoghue and Rabin (2001), DellaVigna and Malmendier (2004), or Heidhues and Koszegi (2010)). In these models, the imperfectly sophisticated agents are often described as being naïve about their time inconsistency, and these agents tend to hold biased beliefs about their future behavior. Thus, time inconsistency induces overconfidence about future behavior. For instance, in DellaVigna and Malmendier (2004), a time inconsistent agent will overestimate the probability of, at some point in the near future paying a cost in order to earn a larger benefit in the distant future for an overall net benefit. They refer to this type of good as an *investment good* and show that (i) naïve consumers will under-consume these types of goods and (ii) firms will alter equilibrium contracts to exploit this naivete. These results have the same flavor of those where a principal contracts with an overconfident agent, so models with time-inconsistent agents might arguably warrant inclusion in a paper like this one.

However, time inconsistency is omitted for a few reasons. First, the formal definition of time inconsistency based on quasi-hyperbolic discounting is relatively more stable across models (although there are exceptions, such as Eliaz and Spiegler (2006)) so that it is at least closer to common knowledge what is meant by time inconsistency without having to state a full-fledged model. Second, that definition of quasi-hyperbolic discounting, as a difference in the relative discount factor between today and tomorrow as compared to two consecutive periods in the future, is applicable to any dynamic model that involves discounting, so it does not suffer from the same drawbacks some definitions of overconfidence have where they are compatible with some models but not others. Finally, the framework presented above is not well suited to an extension that allows for introducing quasi-hyperbolic discounting, and it would severely complicate the analysis. As such, I leave a review of contracting with time-inconsistent agents for future work.

3.4.2 Conclusion

In this paper, I have taken the first step in systematically analyzing the various forms of overconfidence found in the behavioral contract theory literature. By applying them all to a standard information gathering contracting model, I have shown that the effects overconfidence has on the agent's decision-making and the principal's payoff can be sensitive to the precise definition of overconfidence being used. Specifically, overconfidence can either improve or harm incentives for the agent to participate, to gather information, either, or both, and increases in overconfidence can either increase or decrease the principal's profit.

This paper serves partially as a review of the literature on contracts for the overconfident, but its results also have implications for future work in this area. First, this paper is, to my knowledge, the only one to apply different definitions of overconfidence to the same model in order to identify variations in the model's predictions. In fact, alternative definitions of overconfidence are rarely even mentioned (see De La Rosa (2011) or Grubb (2015) for exceptions). Since the extent to which a model's predictions may vary with the definition of overconfidence it uses is now understood, future research should make an effort to either justify whichever definition it proposes, or to discuss the robustness of its predictions to changes in way overconfidence is modeled.

Second, empirical work on overconfidence has primarily focused on identifying a specific type of overconfidence in a particular context, but this paper's results demonstrate that it is possible to distinguish between one type of overconfidence and another, i.e., that different definitions of overconfidence yield different testable implications. To my knowledge, there is only one paper that has explored this idea, Herz et al (2014), which studies the differing effects of overoptimism and overprecision on innovation in an experimental setting. However, there are many other ways that overconfidence can influence decisions-making, as this paper demonstrates. There are still considerable opportunities to explore the ways various types of overconfidence influence any number of settings.

Bibliography

Aktas, N., Louca, C., and Petmezas, D. (2019). CEO overconfidence and the value of corporate cash holdings. *Journal of Corporate Finance* 54, 85-106.

Barber, B. and Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance* 55 (2), 773-806.

Ben-David, I., Graham, J., and Harvey, H. (2013). Managerial miscalibration. *The Quarterly Journal of Economics*, 1547-1584.

Benabou, R. and Tirole, J. (2016). Bonus culture: competitive pay, screening, and multitasking. *Journal of Political Economy* 124 (2), 305-370.

Biais, B., Hilton, D., Mazurier, K., and Pouget, S. (2005). Judgmental overconfidence, self-monitoring, and trading performance in an experimental financial market. *The Review of Economic Studies* 72, 287-312.

Bond, P. and Newman, A. (2009). Prohibitions on punishments in private contracts. *Journal of Financial Intermediation* 18, 526-540.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *The American Economic Review* 89 (1), 306-318.

Crémer, J. and Khalil, F. (1992). Gather information before signing a contract. *The American Economic Review* 82 (3), 566-578.

Crémer, J. and Khalil, F. (1994). Gathering information before the contract is offered: the case of two states of nature. *European Economic Review* 38, 675-782.

Crémer, J., Khalil, F., Rochet, J. (1998a). Contracts and productive information gathering. *Games and Economic Behavior* 25, 174-193.

Crémer, J., Khalil, F., Rochet, J. (1998b). Strategic information gathering before a contract is offered. *Journal of Economic Theory* 81, 163-200.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *The American Economic Review* 89 (1), 306-318.

Compte, O. and Jehiel, P. (2008). Gathering information before signing a contract: a screening perspective. *The International Journal of Industrial Organization* 26, 206-212.

Dai, C., Lewis, T., and Lopomo, G. (2006). Delegating management to experts. *The RAND Journal of Economics* 37 (3), 503-520.

De La Rosa, L. (2011). Overconfidence and moral hazard. *Games and Economic Behavior* 73, 429-451.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature* 47 (2), 315-372.

DellaVigna, S. and Malmendier, U. (2004). Contract design and self-control: Theory and evidence. *The Quarterly Journal of Economics* 2, 353-402.

DellaVigna, S., and Malmendier, U. (2006). Paying not to go to the gym. *The American Economic Review* 96 (3), 694-719.

Downs, J. (2019). Information gathering by overconfident agents. *Working paper.*

Downs, J. (2020a). Information gathering by overconfident agents. *Working Paper (R&R Journal of Economics and Management Strategy.* Chapter 1 of this dissertation.

Downs, J. (2020b). Overconfidence, competition, and market efficiency. *Working Paper*. Chapter 2 of this dissertation.

Eliaz, K. and Spiegler, R. (2006). Contracting with diversely naïve agents. *The Review of Economic Studies* 73 (3). 689-714.

Eliaz, K. and Spiegler, R. (2007). A mechanism-design approach to speculative trade. *Econometrica* 75 (3), 875-884.

Eliaz, K. and Spiegler, R. (2008a). Optimal speculative trade among large traders. *Review of Economic Design* 12, 45-74.

Eliaz, K. and Spiegler, R. (2008b). Consumer optimism and price discrimination. *Theoretical Economics* 3, 459-497.

Fang, H., and Moscarini, G. (2005). Morale hazard. *Journal of Monetary Economics* 52, 749-777.

Finkle, A. (2005). Relying on information acquired by a principal. *International Journal of Industrial Organization* 23, 263-278.

Gabaix, X., and Laibson, D., (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics* 121 (2), 505-540.

Gervais, S. and Goldstein, I. (2007). The Positive effects of biased self-perception in firms. *Review of Finance* 11, 453-496.

Gervais, S., Heaton, J. B., and Odean, T. (2011). Overconfidence, compensation contracts, and capital budgeting. *The Journal of Finance* 66 (5), 1735-1777.

Goel, A. and Thakor, A. (2008). Overconfidence, CEO selection and corporate governance. *The Journal of Finance* 64 (6), 2737-2784.

Grubb, M. (2009). Selling to overconfident consumers. *American Economic Review* 99 (5), 1770-1807.

Grubb, M. (2015). Behavioral consumers in industrial organization: An overview. *Review of Industrial Organization* (2015) 27, 247-258.

Grubb, M. and Osborne, M. (2015). Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review* 105 (1), 234-271.

He, Y., Chen, C., and Hu, Y. (2019). Managerial overconfidence, internal financing, and investment: Evidence from China. *Research in International Business and Finance* 47, 501-510.

Heidhues, P. and Köszegi, B. (2010). Exploiting naivete about self-control in the credit market. *American Economic Review* 100 (5), 2279-2303

Heidhues, P. and Köszegi, B. (2015). On the welfare costs of naivete in the US credit-card market. *Review of Industrial Organization* 47, 341-354.

Herwig, F. and Müller, D. (2016). Overconfidence in the market for lemons. *Scandanavian Journal of Economics* 118 (2), 354-371.

Herz, H., Schunk, D., and Zehnder, C. (2014). How do judgmental overconfidence and overoptimism shape innovative activity? *Games and Economic Behavior* 83, 1-23.

Hoppe, E. and Schmitz, P. (2010). The costs and benefits of additional information in agency models with endogenous information structures. *Economic Letters* 107, 58-62.

Hoppe, E. and Schmitz, P. (2013). Public-private partnerships versus traditional procurement: innovation incentives and information gathering. *The RAND Journal of Economics* 44 (1), 56-74.

Iossa, E. and Martimort, D. (2015). Pessimistic information gathering. *Games and Economic Behavior* 91, 75-96.

Kahneman, D., Wakker, P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics* 112 (2) 375-405.

Kessler, A. (1998). The value of ignorance. *The RAND Journal of Economics* 29 (2), 339-354.

Khalil, F., Kim, D. and Shin, D. (2006). Optimal task design: to integrate or separate planning and implementation? *Journal of Economics and Management Strategy* 15 (2), 457-478.

Kim, J. (2015). Managerial beliefs and incentive policies. *Journal of Economic Behavior and Organization* 119, 84-95.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics* 112 (2), 443-478.

Landier, A. and Thesmar, D. (2009). Financial contracting with optimistic entrepreneurs. *The Review of Financial Studies* 22 (1), 118-150.

Larkin, I. and Leider, S. (2012). Incentive schemes, sorting, and behavioral biases of employees: Experimental Evidence. *American Economic Journal: Microeconomics* 4 (2), 182-214.

Lewis, T. and Sappington, D. (1997). Information management in incentive problems. *Journal of Political Economy* 106 (4), 796-821.

Luz, V. (2017). Characterization and uniqueness of equilibrium in competitive insurance. *Theoretical Economics* 12, 1349-1391.

Malmendier, U. and Tate, G. (2005). CEO Overconfidence and corporate investment. *The Journal of Finance* 60 (6), 2661-2700.

Malmendier, U. and Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics* 89, 20-43.

Manove, M., and Padilla, J., (1999). Banking (conservatively) with optimists. *The RAND Journal of Economics* 30 (2), 324-350.

Malmedier, U. and Taylor, T. (2015). On the verges of overconfidence. *Journal of Economic Perspectives* 29 (4), 3-8.

O'Donoghue, T. and Rabin, M. (2001). Choice and procrastination. *The Quarterly Journal of Economics* 116 (1), 121-160.

Park, Y. J., and Santos-Pinto, L. (2010). Overconfidence in tournaments: evidence from the field. *Theory and Decision* 69, 143-166.

Pikulina, E., Renneboog, L., and Tobler, P. (2017). Overconfidence and investment: An experimental approach. *Journal of Corporate Finance* 43, 175-192.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *The Quarterly Journal of Economics* 90 (4), 629-649.

Sandroni, A. and Squintani, F. (2007) Overconfidence, insurance, and paternalism. *The American Economic Review* 97 (5), 1994-2004.

Sandroni, A. and Squintani, F. (2013) Overconfidence and asymmetric information: The case of insurance. *Journal of Economic Behavior and Organization* (93), 149-165.

Schmitz, P. (2008). Information gathering and the hold-up problem in a complete contracting framework. *Economic Letters* 101, 268-271

Santos-Pinto, L. (2008). Positive self-image and incentives in organizations. *The Economic Journal* 118, 1315-1332.

Santos-Pinto, L. (2012). Labor market signaling and self-confidence: wage compression and the gender pay gap. *Journal of Labor Economics* 30 (4), 873-913.

Sautmann, A. (2013). Contracts for agents with biased beliefs: Some theory and an experiment. *American Economic Journal: Microeconomics* 5 (3), 124-156.

Shin. D., and Yun, S. (2008). Informed principal and information gathering agent. *Review of Economic Design* 12, 229-244.

Szalay, D. (2009). Contracts with endogenous information. *Games and Economic Behavior* 65, 586-625.

Terstiege, S. (2012). Endogenous information and stochastic contracts. *Games and Economic Behavior* 76, 535-547.

Terstiege, S. (2016). Gathering imperfect information before signing a contract. *Games and Economic Behavior* 97, 70-87.

Van Den Steen, E. (2004). Rational overoptimisim (and other biases). *American Economic Review* 94(4), 1141-1151.

Appendix A

Technical Details, Chapter 1

Appendix A.1: Proofs from Section 1.2

Proposition 1: The first part follows from the fact that the set of transfers satisfying both incentive constraints is non-empty and (A1), which implies that the optimal contract that induces information gathering yields the principal a higher profit than any feasible contract that does not. The second part follows from comparing the principal's profit under agreement to that of overconfidence:

$$\Pi_0(\eta) = \eta[V - \beta_L] < \eta[V - \beta_L] + (\tilde{\eta} - \eta)\Delta\beta = \Pi_0(\tilde{\eta}) \Leftrightarrow \eta < \tilde{\eta}$$

Proposition 2: Let η^* be the larger root of the equation $c = \eta^*(1 - \eta^*)\Delta\beta$. By (A2), $\eta^* > \eta$ so (C1) is satisfied for all $\tilde{\eta} \in (\eta, \eta^*]$. When this is the case, the set of contracts satisfying all incentive constraints is non-empty, which, along with (A1), demonstrates the first part. Since η^* is the larger root, $\eta^*(1 - \eta^*)\Delta\beta$ is decreasing in η^* at η^*, so any $\tilde{\eta} > \eta^*$ violates (C1), which demonstrates the second part.

Proposition 3: Define $\tilde{\eta}_x$ as the value of $\tilde{\eta}$ that satisfies:

$$\Pi_1(\tilde{\eta}) = \eta[V - \beta_L] + (\tilde{\eta} - \eta)\Delta\beta - \frac{1 - \eta}{1 - \tilde{\eta}}c = V - \beta_A = \Pi_u(\tilde{\eta})$$

This value is:

$$\tilde{\eta}_x = \frac{V - \beta_H}{V - \beta_H + c} \in (0,1)$$

Straightforward calculation shows that (C1) is satisfied at $\tilde{\eta}_x$, so information gathering is indeed feasible at $\tilde{\eta}_x$. Any $\tilde{\eta} > \tilde{\eta}_x$ sets $\Pi_u(\tilde{\eta}) > \Pi_1(\tilde{\eta})$, which demonstrates the first part. Define $\tilde{\eta}_{max}$ as the value of $\tilde{\eta}$ that maximizes $\Pi_1(\tilde{\eta})$, which satisfies:

$$\frac{d\Pi_1(\tilde{\eta})}{d\tilde{\eta}} = \Delta\beta - \frac{1-\eta}{(1-\tilde{\eta})^2}c = 0$$

This expression makes it clear that $\tilde{\eta}_{max}$ constitutes a global maximum, since $\frac{d\Pi_1(\tilde{\eta})}{d\tilde{\eta}}$ is strictly decreasing in $\tilde{\eta}$. The value $\tilde{\eta}_{max}$ is:

$$\tilde{\eta}_{max} = 1 - \left(\frac{(1-\eta)c}{\Delta\beta}\right)^{\frac{1}{2}}$$

The second part can be demonstrated by proving that for some $c > 0$ we have:

$$\tilde{\eta}_{max} < \tilde{\eta}_x$$

After substitution and some rearranging, this inequality becomes:

$$\Delta\beta c < (1-\eta)(V - \beta_H + c)^2$$

Since $\Delta\beta c \to 0$ as $c \to 0$ and the right-hand side is strictly positive for any $0 \le c < -(V - \beta_H)$, there exists some $\tilde{c}$ small enough such that the inequality $\tilde{\eta}_{max} < \tilde{\eta}_x$ holds for all $c < \tilde{c}$.

Appendix A.2: Proofs from Section 1.3

Proposition 2.A.: Let η^* be the smaller root of $c = \eta^*(1-\eta^*)\Delta\beta$. By (A2), $\eta^* < \eta$ so (C1) is satisfied for all $\tilde{\eta} \in [\eta^*, \eta)$. When this is the case, the set of contracts satisfying all incentive constraints is non-empty, which, along with (A1), demonstrates the first part. Since η^* is the smaller root, $\eta^*(1-\eta^*)\Delta\beta$ is increasing in η^* at η^*, so any $\tilde{\eta} < \eta^*$ violates (C1), which demonstrates the second part.

Proposition 4: By definition, η^{**} is the value of $\tilde{\eta}$ sets $V - \tilde{\eta}\beta_L - (1-\tilde{\eta})\beta_H = 0$. Since the left-hand side of that equation is increasing in $\tilde{\eta}$, we have:

$$V - \beta_A > 0 \Leftrightarrow \tilde{\eta} > \eta^{**}$$

and vice versa. The results follow.

Proposition 3.A.: For the first part, let

$$\Pi_1(\tilde{\eta}) = \eta[V - \beta_L] - \frac{\eta}{\tilde{\eta}}c$$

$$\Pi_u(\tilde{\eta}) = V - \beta_A$$

By (A1) and the assumption that $V - \beta_P > 0$ from Section 2.c., we have that

$$\Pi_1(\eta) > \Pi_u(\eta) > 0$$

Since $\Pi_1(\tilde{\eta}) \to -\infty$ as $\tilde{\eta} \to 0$ and $V - \beta_H \in (-\infty, 0)$, it is clear that there exists some $\tilde{\eta}_x < \eta$ such that $\Pi_1(\tilde{\eta}) < \max\{\Pi_u(\tilde{\eta}), 0\}$ whenever $\tilde{\eta} < \tilde{\eta}_x$. Let η^* be the smaller root of $\eta^*(1-\eta^*)\Delta\beta$. If $\eta^* < \tilde{\eta}_x$, then for all $\tilde{\eta} \in [\eta^*, \tilde{\eta}_x)$, information gathering is feasible but is not induced in the optimal contract and the statement follows. Additionally, the principal's value function for the case of diffidence is

$$\Pi(\tilde{\eta}) = \mathbb{1}_{(\tilde{\eta}\in[\eta^*,\eta])} \max\{\Pi_1(\tilde{\eta}), \Pi_u(\tilde{\eta}), 0\} + \mathbb{1}_{(\tilde{\eta}<\eta^*)} \max\{\Pi_u(\tilde{\eta}), 0\}$$

Since $\Pi_1(\tilde{\eta})$ and $\Pi_u(\tilde{\eta})$ are strictly increasing in $\tilde{\eta}$, $\Pi(\tilde{\eta})$ is non-decreasing in $\tilde{\eta}$, which proves the second part.

Appendix B

Technical Details, Chapter 2

Appendix B.1: Proofs for Sections 2.2-2.5

Proof of Proposition 1: After substitution, the objective is:

$$\max_{e_i} q\left[A(\theta_L + e_L) - \frac{1}{2}e_L^2 + \gamma_L e_L\right] + (1-q)\left[A(\theta_H + e_H) - \frac{1}{2}e_H^2 + \gamma_H e_H\right] - (1-q)\Delta\hat{\theta}e_L$$

The first order conditions are:

$$[e_L]: q[A - e_L + \gamma_L] - (1-q)\Delta\hat{\theta} = 0$$

$$[e_H]: (1-q)[A - e_H + \gamma_H] = 0$$

Rearrangement yields the equilibrium effort levels:

$$e_H^M = A + \gamma_H$$

$$e_L^M = A + \gamma_L - \frac{1-q}{q}\Delta\hat{\theta}$$

Substituting these quantities into the binding constraints yields the equilibrium fixed payments:

$$\alpha_H^M = \alpha_H^{MO} + \left[A + \gamma_L - \frac{1-q}{q}\Delta\hat{\theta}\right]\Delta\hat{\theta} > \alpha_H^{MO}$$

$$\alpha_L^M = \alpha_L^{MO} + \frac{1-q}{q}\Delta\hat{\theta}\left[A + 2\gamma_L + \theta_L - \frac{1-q}{2q}\Delta\hat{\theta}\right] > \alpha_L^{MO}$$

Proof of Proposition 2: When Condition 1 is satisfied, neither incentive constraint binds so the objective is:

$$\max_{e_i} q\left[A(\theta_L + e_L) - \frac{1}{2}e_L^2 + \gamma_L e_L\right] + (1-q)\left[A(\theta_H + e_H) - \frac{1}{2}e_H^2 + \gamma_H e_H\right]$$

The first order conditions yield the equilibrium quantities and substitution into the zero profit conditions $[ZP_i]$ yields the equilibrium fixed payments.

When condition 1 is violated, $[IC_H]$ will be binding. In equilibrium, high types must be offered $(e_H^{CO}, \alpha_H^{CO})$ as any other contract tightens $[IC_H]$ and is less attractive to high types. Then the binding $[IC_H]$ and $[ZP_L]$ determine e_L^{RS}, which is the solution to the following equation:

$$\frac{1}{2}A^2 + A\theta_H + A\gamma_H + \frac{1}{2}\gamma_H^2 = \frac{1}{2}(e_L^{RS})^2 + e_L^{RS}(A + \Delta\theta + \gamma_H) + A\theta_L$$

There are two solutions:

$$e_L^{RS} = A + \Delta\theta + \gamma_H \pm \sqrt{2\Delta\theta\gamma_H + \Delta\theta^2}$$

The larger root can be ruled out as the resulting contract would violate $[IC_L]$.

Proof of Lemma 1: "If": Suppose not: the RS contracts are interim efficient but *not* the solution to P^D. Then there must be another allocation, call it $(\tilde{e}_i, \tilde{\alpha}_i)_{i\in\{L,H\}}$, that satisfies all the constraints and yields strictly positive profit ($[ZP]$ is a strict inequality). In that case, the firm could increase either $\tilde{\alpha}_L$ or $\tilde{\alpha}_H$ by a small amount, keeping $[ZP]$ as a strict inequality and strictly increasing either $U(\tilde{e}_L, \tilde{\alpha}_L; \hat{\theta}_L)$ or $U(\tilde{e}_H, \tilde{\alpha}_H; \hat{\theta}_H)$ to be strictly larger than U_L^{RS} or U_H^{RS}. This new allocation Pareto Dominates the RS contracts, which contradicts the supposition that the RS contracts are interim efficient.

"Only if": Suppose not: the RS contracts are the solution to P^D but *not* interim efficient. Since the RS contracts bind $[ZP]$, this implies there exists another allocation (set of contracts) $(\tilde{e}_i, \tilde{\alpha}_i)_{i\in\{L,H\}}$ such that $U(\tilde{e}_H, \tilde{\alpha}_H; \hat{\theta}_H) \geq U(e_H^{RS}, \alpha_H^{RS}; \hat{\theta}_H)$ and $U(\tilde{e}_L, \tilde{\alpha}_L; \hat{\theta}_L) \geq U(e_L^{RS}, \alpha_L^{RS}; \hat{\theta}_L)$, with at least one strict inequality, and that satisfy $[ZP]$. Suppose, without loss of generality, that $U(\tilde{e}_H, \tilde{\alpha}_H; \hat{\theta}_H) > U(e_H^{RS}, \alpha_H^{RS}; \hat{\theta}_H)$. In that case, the objective in P^D could be increased by reducing $\tilde{\alpha}_H$ without affecting any constraints. This contradicts the supposition that the RS contracts are the solution to problem P^D.

Proof of Proposition 3: When condition 1 is satisfied, no other set of feasible contracts offer either type of worker more utility, so no firm can profitably offer any other menu of contracts. When condition 1 is violated, I show that when $q \leq q_H$, the RS contracts are the solution to problem P^D, which, by Lemma 1, implies they are the unique competitive equilibrium contracts.

First, notice that when $[IR_L], [IR_H], [IC_H]$, and $[ZP]$ are all equalities at the optimum, the solution to P^D is the RS contracts. Now, other than $[ZP]$, problem P^D is a monopolistic screening problem with type dependent outside options, but U_H^{RS} and U_L^{RS} are such that $[IC_L]$ can never bind at the optimum, so it can be ignored. The solution when ignoring $[IC_H]$ is the observable type contracts, which violate $[IC_H]$, so that constraint binds. Then $[IR_L]$ must bind as well, otherwise the objective could be increased with a reduction in α_L. The remaining constraints are $[IR_H]$ and $[ZP]$. If both are ignored, the solution to P^D involves the monopsony effort levels e_L^M, e_H^M. In that case, the contracts satisfy $[ZP]$ with strict inequality (by assumption from the monopsony case that the monopsonist can make a positive profit in equilibrium). Using the binding $[IC_H]$ and $[IR_L]$, the left-hand side of $[IR_H]$ is given by $U_L^{RS} + \Delta\hat{\theta} e_L^M$, so $[IR_H]$ can be written as:

$$\Delta\hat{\theta} e_L^M \geq U_H^{RS} - U_L^{RS} = \Delta\hat{\theta} e_L^{RS}$$

where the equality comes from the fact that the RS contracts bind $[IC_H]$. Therefore, the solution to P^D exactly satisfies $[IR_H]$ with equality whenever $e_L^{RS} \geq e_L^M \Leftrightarrow q \leq q_H$, in which case $[IR_H], [IR_L], [IC_H]$ are all equalities. Finally, $[ZP]$ is also an equality in this case as the RS menu yields workers the most utility among incentive compatible menus holding firm profit from each type at exactly zero. Thus, when $q \leq q_H$, the RS contracts are the solution to P^D.

Proof of Proposition 4: Since $e_H^{\zeta} = A + \gamma_H$ for all $\zeta \in \{M, CO, RS\}$, we have

$$S^{\zeta} = qS_L\left(e_L^{\zeta}\right) + (1-q)S_H(A+\gamma_H)$$

Notice that $S_L(e_L^{\zeta})$ is strictly decreasing at all $e_L > e_L^*$. For part (i), $e_L^{CO} > e_L^{RS} > e_L^* \Rightarrow L^{CO} > L^{RS}$. For part (ii), it has been shown that $e_L^M < e_L^C$. Therefore, a sufficient condition for $L^{CO} > L^M$ is $e_L^M > e_L^*$, which is the case if and only if $q > q_L$. For part (iii), Proposition 3 proves that $e_L^{RS} > e_L^M \Leftrightarrow q < q_H$, so $e_L^M > e_L^*$ is a sufficient condition for $L^{RS} > L^M$, which is shown to be the case when $q > q_L$ when proving part (ii). Finally, we have:

$$q_L \equiv \frac{\Delta\theta + \Delta\gamma}{\Delta\theta + \gamma_H} < \frac{\Delta\theta + \Delta\gamma}{X} \equiv q_H \Leftrightarrow \Delta\theta + \gamma_H > X$$

Squaring both sides:

$$\Delta\theta^2 + \gamma_H^2 + 2\Delta\theta\gamma_H > 2\Delta\theta\gamma_H + \Delta\theta^2 \Leftrightarrow \gamma_H^2$$

The inequality holds for any $\gamma_H \neq 0$.

Proof of Proposition 5: First, note that:

$$\hat{U}^{CO} = q[S_L(e_L^{CO}) + \gamma_L e_L^{CO}] + (1-q)[S_H(e_H^{CO}) + \gamma_H e_H^{CO}]$$

$$\hat{U}^{RS} = q[S_L(e_L^{RS}) + \gamma_L e_L^{RS}] + (1-q)[S_H(e_H^{CO}) + \gamma_H e_H^{CO}]$$

$$\hat{U}^{M} = (1-q)\Delta\hat{\theta}e_L^M$$

Also note that $\hat{U}^{\zeta}$ is strictly increasing in e_L^{ζ} for $e_L^{\zeta} < e_L^{CO}$ and $\zeta \in \{CO, RS\}$. Then in part (i), $\hat{U}^{CO} > \hat{U}^{RS}$ follows from the fact that $e_L^{CO} > e_L^{RS}$. As for $\hat{U}^{RS} > \hat{U}^M$, the (implicit) assumption that the monopsonist profits from employing high-type workers holds when $(1-q)\left[S_H(e_H^M) + \gamma_H e_H^M - \Delta\hat{\theta}e_L^M\right] > 0$. Since $e_L^{CO} = e_L^M$, the second term of $\hat{U}^{RS}$ is greater than $\hat{U}^M$, and since the first term of $\hat{U}^{RS}$ is strictly positive, that $\hat{U}^{RS} > \hat{U}^M$ follows.

For part (ii), $U^{\zeta} = S^{\zeta}$ for $\zeta \in \{CO, RS\}$, so the first inequality follows from Proposition 4. For the second inequality, we have that:

$$U^{C} = \widehat{U}^{C} - q\gamma_L e_L^{CO} - (1-q)\gamma_H e_H^{CO}$$

$$U^{M} = \widehat{U}^{M} - q\gamma_L e_L^{M} - (1-q)\gamma_H e_H^{M}$$

Then, using the expressions for $\widehat{U}^{CO}$ and $\widehat{U}^{M}$ given above:

$$U^{CO} - U^{M} = q[S_L(e_L^{CO}) + \gamma_L e_L^{M}] + (1-q)[S_H(e_H^{CO}) + \gamma_H e_H^{CO} - \Delta\hat{\theta} e_L^{M}]$$

The second term is strictly positive by the same argument used in part (i), and the first term is strictly positive by the assumption that $S_i(e_i)$ is positive in any equilibrium from Section 2.1.

Appendix B.2: Analysis and Proofs for Section 2.6

Analysis for Section 4.1.

There is one case that can be ruled out immediately. We cannot have $\Delta\hat{\theta} < 0, \gamma_H > 0$ and $\gamma_L < 0$ since $\Delta\hat{\theta} < 0$ and $\gamma_H > 0 \Rightarrow \gamma_L > 0$.

Monopsony: The case where $\Delta\hat{\theta} > 0$ is covered in Section 2. When $\Delta\hat{\theta} < 0$, the statement of the monopsonist's problem is unchanged, but which constraints are binding is reversed: $[IC_L]$ and $[IR_H]$ will bind instead. After substitution, the monopsonist's problem becomes:

$$\max_{e_i} q\left[S_L(e_L) + \gamma_L e_L + \Delta\hat{\theta} e_H\right] + (1-q)[S_H(e_H) + \gamma_H e_H]$$

Rearranging the FOCS yields:

$$e_L^{M} = A + \gamma_L$$

$$e_H^{M} = A + \gamma_H + \frac{q}{1-q}\Delta\hat{\theta} < A + \gamma_H = e_H^{CO}$$

The downward distortion in e_H^{M} improves welfare if $\gamma_H > 0$ and $q < \frac{\gamma_H}{\gamma_H - \Delta\hat{\theta}} \in (0,1)$.

Perfect Competition: Deriving the RS contracts with general differences in beliefs follows the same logic of that in Section 2. First, determine if either Condition 1 or 2 are violated. If not, we have $e_i^{RS} = e_i^{CO}$. If, for example, Condition 1 is violated, then $[IC_H]$ binds, so $e_H^{RS} = e_H^{CO}$ and e_L^{RS} will have to be distorted to satisfy $[IC_H]$. The direction of this distortion depends on the sign of $\Delta\hat{\theta}$. If $\Delta\hat{\theta} > 0$, then the smallest possible (utility maximizing) incentive compatible distortion is downward, $e_L^{RS} < e_L^{CO}$.[60] If $\Delta\hat{\theta} < 0$, then the smallest possible incentive compatible distortion is upward, $e_L^{RS} > e_L^{CO}$. Whether that distortion improves efficiency (as compared e_L^{CO}) depends on the sign of γ_L. If $\gamma_L > 0$, then $\Delta\hat{\theta} > 0$ results in the welfare improving distortion $e_L^{RS} < e_L^{CO}$. If $\gamma_L < 0$, then $\Delta\hat{\theta} < 0$ results in the welfare improving distortion $e_L^{RS} > e_L^{CO}$.

Note that only one incentive constraint can bind in a competitive equilibrium (so that checking Conditions 1 and 2 to determine which incentive constraints bind is a valid approach). This follows from Lemma 1: competitive equilibrium contracts must be the solution to problem P^D, but the only way for both incentive constraints to be satisfied with an equality in that problem is in a pooling equilibrium. Following the result in Rothschild and Stiglitz (1976), no such pooling equilibrium exists since a firm could profitably deviate from such an equilibrium by offering a contract that only attracts the more efficient type of worker (according to the workers' beliefs) and yields a strictly positive profit.

The steps described above can be used to derive the RS contracts for each possible case, but it must also be shown that they constitute the competitive equilibrium contracts. Several cases can be ruled out: those where different incentive constraints bind in the RS equilibrium and

[60] In a model with more general cost and revenue functions, solving for the value of e_L^{RS} using the binding $[IC_H]$ might involve several solutions. BT put conditions on those functions that limit the number of solutions to two and use the same argument about which solution should be used, i.e., the LCS or least cost separating allocation of effort.

problem P^D from Lemma 1. For example, suppose that the RS contracts bind $[IC_H]$, while $\Delta\hat{\theta} < 0$ so that $[IC_L]$ is the binding constraint in problem P^D. Then we have:

$$U_H^{RS} = U_L^{RS} + \Delta\hat{\theta} e_L^{RS}$$

$$U_H^D = U_L^D + \Delta\hat{\theta} e_H^D$$

The superscript D indicates values at the solution to P^D. Since $[IC_L]$ binds in P^D, so must $[IR_H]$, which gives us that $U_H^D = U_H^{RS}$. Then, $[IR_L]$ also binds when:

$$U_H^{RS} - \Delta\hat{\theta} e_L^{RS} > U_H^{RS} - \Delta\hat{\theta} e_H^D \Leftrightarrow e_H^D > e_L^{RS}$$

However, since $\Delta\hat{\theta} < 0$ and Condition 1 is violated, $e_L^{RS} > e_L^{CO}$, so $[IR_L]$ binding requires $e_H^D > e_L^{CO} > e_H^{CO}$, but $e_H^D < e_H^{CO} < e_L^{CO}$ in the solution to P^D. Thus, the RS contracts are not interim efficient if either Condition 1 is violated and $\Delta\hat{\theta} < 0$ or Condition 2 is violated and $\Delta\hat{\theta} > 0$.[61]

Cases that remain are when both Condition 1 and 2 are satisfied, when Condition 1 is violated and $\Delta\hat{\theta} > 0$, and when Condition 2 is violated and $\Delta\hat{\theta} < 0$. When both conditions are satisfied, the RS contracts are also the observable type contracts, which are straightforwardly a competitive equilibrium. The case where Condition 1 is violated, $\gamma_i > 0$, and $\Delta\hat{\theta} > 0$ is what is considered in the baseline model. All that's left is the derive conditions under which the RS contracts are a competitive equilibrium when Condition 2 is violated, $\gamma_i < 0$, and $\Delta\hat{\theta} < 0$.

The logic is analogous to that of the baseline model; $[IC_L]$ is binding at the RS contracts and in problem P^D, so $[IR_H]$ also binds in problem P^D. What is needed is for $[IR_L]$ to also bind in problem P^D, which is the case when $e_H^M > e_H^{RS}$. Using the binding $[IC_L]$, we have:

$$e_H^{RS} = (A - \Delta\theta + \gamma_L) - \sqrt{\Delta\theta^2 - 2\gamma_L \Delta\theta}$$

From the monopsony problem, we have:

---

[61] The proof for this second case is analogous to the first.

$$e_H^M = A + \gamma_H + \frac{q}{1-q}\Delta\hat{\theta}$$

Thus, the RS contracts are interim efficient if:

$$\gamma_H + \frac{q}{1-q}\Delta\hat{\theta} > \gamma_L - \Delta\theta - \sqrt{\Delta\theta^2 - 2\gamma_L\Delta\theta}$$

$$\Leftrightarrow q < q_H' \equiv \frac{\sqrt{\Delta\theta^2 - 2\gamma_L\Delta\theta} + \Delta\hat{\theta}}{\sqrt{\Delta\theta^2 - 2\gamma_L\Delta\theta}} \in (0,1)$$

Finally, since $\gamma_H < 0$, monopsony is more efficient than the RS allocation whenever $e_H^{RS} < e_H^M < e_H^*$, but since $e_H^M < e_H^*$ for any $q > 0$, $e_H^M > e_H^{RS}$ is sufficient for this to be the case.

Optimal Contracts for Section 4.2.

When types are observable and workers are overconfident, optimal contracts in either monopsony or perfect competition maximize, for each type:

$S_i(e_i) \equiv Ae_i - \frac{\hat{\theta}_i}{2}e_i^2 \Rightarrow e_i^{CO} = \frac{A}{\hat{\theta}_i}$.

The monopsonist's problem is:

$$P^M: \max_{\mathrm{e_i},\alpha_i} q[Ae_H - t_H] + (1-q)[Ae_L - t_L]$$

$$[IR_L]: t_L - \frac{\hat{\theta}_L}{2}e_L^2 \geq 0$$

$$[IR_H]: t_H - \frac{\hat{\theta}_H}{2}e_H^2 \geq 0$$

$$[IC_L]: t_L - \frac{\hat{\theta}_L}{2}e_L^2 \geq t_H - \frac{\hat{\theta}_L}{2}e_H^2$$

$$[IC_H]: t_H - \frac{\hat{\theta}_H}{2}e_H^2 \geq t_L - \frac{\hat{\theta}_H}{2}e_L^2$$

With $\Delta\hat{\theta} > 0$, it is straightforward to show that the solution involves $[IC_L], [IR_H]$ binding and the other two constraints slack. Substitution yields the simplified problem:

$$P^M\colon \max q\left[Ae_H-\frac{\hat{\theta}_H}{2}e_H^2\right]+(1-q)\left[Ae_L-\frac{\hat{\theta}_L}{2}e_L^2-\frac{\Delta\hat{\theta}}{2}e_H^2\right]$$

Straightforward maximization yields the monopsony effort levels:

$$e_L^M=\frac{A}{\hat{\theta}_L}=e_L^{CO}$$

$$e_H^M=\frac{A}{\hat{\theta}_H+\frac{1-q}{q}\Delta\hat{\theta}}<e_H^{CO}$$

In perfect competition with observable types, the equilibrium contracts are the solution to:

$$\max t_i-\frac{\hat{\theta}_i}{2}e_i^2$$

$$[ZP_i]\colon Ae_i-t_i\geq 0$$

The equilibrium contracts are:

$$e_i^{CO}=\frac{A}{\hat{\theta}_i},t_i^{CO}=\frac{A^2}{\hat{\theta}_i}$$

Substitution and some rearrangements show that these contracts are incentive compatible. Define:

$U_i^{CO}\equiv t_i^{CO}-\frac{\hat{\theta}_i}{2}\left(e_i^{CO}\right)^2=\frac{A^2}{2\hat{\theta}_i}$.

Then, incentive compatibility requires:

$$\frac{\Delta\hat{\theta}}{2}(e_H^{CO})2\leq U_L-U_H\leq\frac{\Delta\hat{\theta}}{2}(e_L^{CO})^2$$

After substitution and cancelling common terms, this simplifies to:

$$\frac{1}{\hat{\theta}_H^2}<\frac{1}{\hat{\theta}_H\hat{\theta}_L}<\frac{1}{\hat{\theta}_L^2}$$

These inequalities hold since, since $\hat{\theta}_H>\hat{\theta}_L$. Therefore, the observable type contracts are also the unique competitive equilibrium contracts when the worker's type is not observable.

Proof of Proposition 7: Using the definition of L^{ζ} given in Section 3, the fact that $e_L^{CO} = e_L^M$, and the fact that $S_H(e)$ is strictly decreasing in e for all $e < e_L^{CO}$, a sufficient condition for $L^{CO} < L^M$ is

$$e_H^M > \frac{A}{\theta_H} \Leftrightarrow q > \frac{\Delta\hat{\theta}}{\Delta\hat{\theta}+\gamma_H} \equiv \tilde{q}_L.$$

Appendix C

Technical Details, Chapter 3

Analysis for Sections 3.1-3.8

3.1. Assumptions (A1) and (A2) for 3.1. are:

(A1): $c < \min\{\eta(V-\beta_L), -(1-\eta)(V-\beta_H)\}$

(A2): $c < \eta(1-\eta)\Delta\beta$

Proof of Proposition 1: The principal's problem of inducing information gathering is:

$$\max_{t_H, t_L} \eta(V - t_L) + (1-\eta)\,(-t_H)$$

$$\eta(t_L - \beta_L) + (1-\eta)(t_H) - \hat{c} \geq 0 \qquad [PC]$$

$$\eta(t_L - \beta_L) + (1-\eta)(t_H) - \hat{c} \geq t_L - \beta_A \qquad [IG_L]$$

$$\eta(t_L - \beta_L) + (1-\eta)(t_H) - \hat{c} \geq t_H \qquad [IG_H]$$

$$t_L - \beta_L \geq t_H \qquad [IC_L]$$

$$t_H \geq t_L - \beta_H \qquad [IC_H]$$

First notice that $[IG_L] \Rightarrow [IC_H]$ and $[IG_H] \Rightarrow [IC_L]$, so the $[IC]$s can be ignored. The remaining constraints, $[PC], [IG_L]$, and $[IG_H]$, are all relaxed by decreases in $\hat{c}$, so increasing overconfidence makes the agent more willing to participate and to gather information. Since the participation constraint binds at the optimum, decreases in $\hat{c}$ allow the principal to reduce both transfers, so overconfidence unambiguously benefits the principal.

3.2. Assumptions (A1) and (A2) for 3.2. are:

(A1): $c < \min\{\eta(V-\beta_L), -(1-\eta)(V-\beta_H)\}$

(A2): $c < \eta(1-\eta)\Delta\beta$

Proof of Proposition 2: A full proof for this case can be found in Downs (2020a).

3.3. Assumptions (A1) and (A2) for 3.3. are:

(A1): $c < \min\{\eta(V-\beta_L), -(1-\eta)(V-\beta_H)\}$

(A2): $c < \eta(1-\eta)\Delta\beta$

Proof of Proposition 3: The principal's problem of inducing information gathering is:

$$\max_{t_H,t_L} \eta(V-t_L) + (1-\eta)\,(-t_H)$$

$$\eta\big(t_L-\hat{\beta}_L\big) + (1-\eta)(t_H) - c \geq 0 \qquad [PC]$$

$$\eta\big(t_L-\hat{\beta}_L\big) + (1-\eta)(t_H) - c \geq t_L - \hat{\beta}_A \qquad [IG_L]$$

$$\eta\big(t_L-\hat{\beta}_L\big) + (1-\eta)(t_H) - c \geq t_H \qquad [IG_H]$$

$$t_L - \hat{\beta}_L \geq t_H \qquad [IC_L]$$

$$t_H \geq t_L - \hat{\beta}_H \qquad [IC_H]$$

The $[IG]$s imply the $[IC]$s, so they can be ignored. Since $[PC]$ is binding at the optimum, increases in γ_L reduce $\hat{\beta}_L$, which relaxes the participation constraint, while changes in γ_H have no effect on the agent's willingness to participation, since he does not implement the project when costs are high. Substituting in the binding $[PC]$, the principal's problem is:

$$\max_{t_H} \eta\big(V-\hat{\beta}_L\big) - c$$

$$t_H \geq \frac{c}{1-\eta} - \eta\Delta\hat{\beta} \qquad [IG_L]$$

$$0 \geq t_H \qquad [IG_H]$$

Since increases in $\Delta\hat{\beta}$ relax $[IG_L]$, increases in γ_L and decreases in γ_H also relax this constraint, meaning the agent is more willing to gather information. Thus, increases in γ_L and decreases in γ_H weakly benefit the principal.

3.4. Assumptions (A1) and (A2) for 3.4. are:

(A1): $c < \min\{-\eta(1-r)(V-\beta_L)-(1-\eta)r(V-\beta_H), \eta r(V-\beta_L)+(1-\eta)(1-r)(V-\beta_H\}$

(A2): $c < \eta(1-\eta)(2r-1)\Delta\beta$

Proof of Proposition 4: Assume that, without any loss of generality, it is optimal for the principal to have to agent implement the project when he is uninformed. Then the agent is paid the same transfer t_L when he is uninformed as when his realized cost is low. The principal's problem of inducing information gathering is:

$$\max_{t_H,t_L} \pi[\eta(V-t_L)+(1-\eta)(-t_H)]+(1-\pi)(V-t_L)$$

$$\hat{\pi}[\eta(t_L-\beta_L)+(1-\eta)t_H]+(1-\hat{\pi})(t_L-\beta_A)-c \geq 0 \qquad [PC]$$

$$\hat{\pi}[\eta(t_L-\beta_L)+(1-\eta)t_H]+(1-\hat{\pi})(t_L-\beta_A)-c \geq t_L-\beta_A \qquad [IG_L]$$

$$\hat{\pi}[\eta(t_L-\beta_L)+(1-\eta)t_H]+(1-\hat{\pi})(t_L-\beta_A)-c \geq t_H \qquad [IG_H]$$

$$\hat{\pi}[\eta(t_L-\beta_L)+(1-\eta)t_H]+(1-\hat{\pi})(t_L-\beta_A)-c \geq t_L-\beta_A \qquad [IG_U]$$

$$t_L-\beta_L \geq t_H \qquad [IC_{\{LH\}}]$$

$$t_H \geq t_L-\beta_H \qquad [IC_{\{HL\}}]$$

$$t_H \geq t_L-\beta_H \qquad [IC_{\{HU\}}]$$

$$t_L-\beta_A \geq t_H \qquad [IC_{\{UH\}}]$$

The constraints $[IC_{\{LU\}}]$ and $[IC_{\{UL\}}]$ have been omitted since they are tautological. Additionally, several constraints can be ignored: $[IC_{\{UH\}}] \Rightarrow [IC_{\{LH\}}]$, $[IC_{\{HL\}}] \Rightarrow [IC_{\{HU\}}]$, $[IG_L] \Leftrightarrow [IG_U]$, and $[IC_{\{UH\}}]$ and $[IG_L] \Rightarrow [IG_H]$, so all the implied constraints can be ignored. The participation constraint is binding since otherwise the principal could reduce all transfers without violating any constraints and increase her profit. The binding $[PC]$ implies that $U(\emptyset) \equiv t_L - \beta_A \leq 0$ and $t_H \leq 0$, so it must be that $U(\sigma) \equiv \eta(t_L - \beta_L) + (1-\eta)t_H > 0$. Thus, increases in $\hat{\pi}$ relax $[PC]$. Then rewrite $[IG_L]$ as:

$$\hat{\pi}[\eta(t_L - \beta_L) + (1-\eta)t_H - (t_L - \beta_A)] \equiv \hat{\pi}\big(U(\sigma) - U(\emptyset)\big) \geq c$$

The term in the brackets on the LHS is strictly positive meaning increases in $\hat{\pi}$ relax $[IG_L]$. All constraints are relaxed by the agent's overconfidence, meaning the principal's profit is increased by overconfidence.

3.5. Assumptions (A1) and (A2) for 3.5. are:

(A1): $c < \min\{(p+\pi)\eta(V-\beta_L) + (1-p-\pi)(V-\beta_A), -(p+\pi)(1-\eta)(V-\beta_H)\}$

(A2): $c < (p+\pi)\eta(1-\eta)\Delta\beta$

Proof of Proposition 5: The principal's problem of inducing information gathering is:

$$\max_{t_H, t_L} \eta[r(V-t_L) + (1-r)(-t_H)] + (1-\eta)[r(-t_H) + (1-r)(V-t_L)]$$

$$\eta[\hat{r}(t_L-\beta_L) + (1-\hat{r})(t_H)] + (1-\eta)[\hat{r}(t_H) + (1-\hat{r})(t_L-\beta_H)] - c \geq 0 \qquad [PC]$$

$$\eta[\hat{r}(t_L-\beta_L) + (1-\hat{r})(t_H)] + (1-\eta)[\hat{r}(t_H) + (1-\hat{r})(t_L-\beta_H)] - c \geq t_L - \beta_A \qquad [IG_L]$$

$$\eta[\hat{r}(t_L-\beta_L) + (1-\hat{r})(t_H)] + (1-\eta)[\hat{r}(t_H) + (1-\hat{r})(t_L-\beta_H)] - c \geq t_H \qquad [IG_H]$$

$$t_L - \hat{r}\beta_L - (1-\hat{r})\beta_H \geq t_H \qquad [IC_L]$$

$$t_H \geq t_L - \hat{r}\beta_H - (1-\hat{r})\beta_L \qquad [IC_H]$$

Notice that, since $\beta_H > \beta_L$, increases in $\hat{r}$ relax both $[IC_L]$ and $[IC_H]$. The information gathering constraints imply that $t_H \leq 0$ and $t_L - \beta_A \leq 0$, which together with the binding $[PC]$ imply that $t_L - \beta_L > 0$. Finally, $[IC_H]$ implies that $t_H \geq t_L - \beta_H$. Then, increases in $\hat{r}$ relax $[PC], [IG_L]$ and $[IG_H]$ as well. Thus, all constraints are relaxed by increases in $\hat{r}$, implying the results.

3.6. Assumptions (A1) and (A2) for 3.6. are:

(A1): $c < \min\{(p+\pi)\eta(V-\beta_L) + (1-p-\pi)(V-\beta_A), -(p+\pi)(1-\eta)(V-\beta_H)\}$

(A2): $c < (p+\pi)\eta(1-\eta)\Delta\beta$

Proof of Proposition 6: I follow the same logic of the proof of Proposition 4. Assume that it is optimal for the principal to have an uninformed agent implement the project, so that an uninformed agent's transfer is $t_u = t_L$. The principal's problem of inducing information gathering is:

$$\max_{t_H, t_L}(p+\pi)[\eta(V-t_L) + (1-\eta)(-t_H)] + (1-p-\pi)(V-t_L)$$

$$(\hat{p}+\pi)[\eta(t_L-\beta_L) + (1-\eta)t_H] + (1-(\hat{p}+\pi))(t_L-\beta_A) - c \geq 0 \qquad [PC]$$

$$(\hat{p}+\pi)[\eta(t_L-\beta_L) + (1-\eta)t_H] + \left(1-(\hat{p}+\pi)\right)(t_L-\beta_A) - c \geq$$
$$\hat{p}[\eta(t_L-\beta_L) + (1-\eta)(t_H)] + (1-\hat{p})(t_L-\beta_A) \qquad [IG_L]$$

$$(\hat{p}+\pi)[\eta(t_L-\beta_L) + (1-\eta)t_H] + \left(1-(\hat{p}+\pi)\right)(t_L-\beta_A) - c \geq$$
$$\hat{p}[\eta(t_L-\beta_L) + (1-\eta)(t_H)] + (1-\hat{p})(t_H) \qquad [IG_H]$$

$$(\hat{p}+\pi)[\eta(t_L-\beta_L) + (1-\eta)t_H] + \left(1-(\hat{p}+\pi)\right)(t_L-\beta_A) - c \geq$$
$$\hat{p}[\eta(t_L-\beta_L) + (1-\eta)(t_H)] + (1-\hat{p})(t_L-\beta_A) \qquad [IG_U]$$

$$t_L - \beta_L \geq t_H \qquad [IC_{\{LH\}}]$$

$$t_H \geq t_L - \beta_H \qquad [IC_{\{HL\}}]$$

$$t_H \geq t_L - \beta_H \qquad [IC_{\{HU\}}]$$

$$t_L - \beta_A \geq t_H \qquad [IC_{\{UH\}}]$$

As for the agent's incentive to participate, argument is identical to that for Proposition 4 with $\hat{\pi}$ replaced by $\hat{p} + \pi$: the agent is more willing to participate when overconfident. For information gathering incentives, using the same definitions for $U(\sigma)$ and $U(\emptyset)$ as in the proof of Proposition 4, $[IG_L]$ can be written as:

$$\pi\big(U(\sigma) - U(\emptyset)\big) \geq c$$

Clearly, the agent's overestimation of p as $\hat{p}$ has no effect on his incentive to gather information. Since the agent's willingness to gather information is unchanged and he is more willing to participate, the principal unambiguously benefits from the agent's overconfidence.

3.7. Assumptions (A1) and (A2) for 3.7. are:

(A1): $c < \min\{\eta(V - \beta_L), -(1-\eta)(V - \beta_H)\}$

(A2): $c < \eta(1-\eta)\Delta\beta$

Proof of Proposition 7: The principal's problem of inducing information gathering is:

$$\max_{t_H, t_L} (\eta + (1-\eta)\mu)(V - t_L) + (1-\eta)(1-\mu)(-t_H)$$

$$\eta(t_L - \beta_L) + (1-\eta)(t_H) - c \geq 0 \qquad [PC]$$

$$\eta(t_L - \beta_L) + (1-\eta)(t_H) - c \geq t_L - \beta_A \qquad [IG_L]$$

$$\eta(t_L - \beta_L) + (1-\eta)(t_H) - c \geq t_H \qquad [IG_H]$$

$$t_L - \beta_L \geq t_H \qquad [IC_L]$$

$$t_H \geq t_L - \beta_H \qquad [IC_H]$$

Inspection of the constraints demonstrates that the agent's overconfidence, parameterized here by $\mu > 0$, has no effect on the agent's willingness to participation, gather, or reveal information. The only change is in the principal's objective: she encounters an agent who believes his implementation cost is low with probability $\eta + (1-\eta)\mu > \eta$. Since the $[IG]$s imply the $[IC]$s, assumption (A2) implies that there is a non-empty set of transfer pairs (t_L, t_H) that satisfy $[PC]$ and the $[IG]$s, and all yield the principal her maximized profit. One of these pairs is $t_L = \beta_L + c$ and $t_H = c$. Then $V - t_L = V - \beta_L - c > -c = t_H$, which implies that that objective function is strictly increasing in μ. Therefore, the agent's incentives are unaffected by his overconfidence, but overconfidence benefits the principal, nonetheless.

3.8. Assumptions (A1) and (A2) for 3.8. are:

(A1): $c < \min\{\eta(V-\beta_L), -(1-\eta)(V-\beta_H)\}$

(A2): $c < \eta(1-\eta)\Delta\beta$

Proof of Proposition 8: The principal's problem of inducing information gathering is:

$$\max_{t_H, t_L} \eta(V - t_L) + (1-\eta)\,(-t_H)$$

$$\eta\left(t_L - \hat{\beta}_L\right) + (1-\eta)(t_H) - c \geq 0 \qquad [PC]$$

$$\eta\left(t_L - \hat{\beta}_L\right) + (1-\eta)(t_H) - c \geq t_L - \beta_P \qquad [IG_L]$$

$$\eta\left(t_L - \hat{\beta}_L\right) + (1-\eta)(t_H) - c \geq t_H \qquad [IG_H]$$

$$t_L - \hat{\beta}_L \geq t_H \qquad [IC_L]$$

$$t_H \geq t_L - \hat{\beta}_H \qquad [IC_H]$$

The $[IG]s$ imply the $[IC]s$ so they can be ignored. The left hand side of $[PC]$, $[IG_L]$, and $[IG_H]$ are all reduced by the agent's overconfidence since he believes that $\hat{\beta}_L > \beta_L$, while the right-hand sides are unaffected. Since $[PC]$ must bind at the optimum, increases in the agent's overconfidence (increases in $\hat{\beta}_L$) must decrease the principal's profit.