Competitive Pay and Excessive Manager Risk-Taking

Jen-Wen Chang,* Simpson Zhang†

February 17, 2018

Abstract
Since the financial crisis, researchers have debated whether executive compensation drove excessive risk-taking or corporate executives simply underestimated the risks of various investments. Through a principal-agent model with heterogeneous beliefs, we show that principals offer contracts that incentivize safe behavior when competition for managerial talent is low. However, intense competition results in contracts that incentivize risk-taking. We show that factors that increase the intensity of competition include greater search efficiency, larger project scales, and higher debt funding, all of which may be prevalent during a financial bubble.

JEL Codes: D86, G38, M12

Keywords: competition, compensation contracts, overoptimism

1 Introduction
The history of financial crises, including the recent subprime loan crisis, is a history of bubbles bursting: a buildup of financial valuations, fueled by overly-optimistic beliefs and

---

*California State University, Fullerton. jenwenc@gmail.com

†Office of Financial Research. Views and opinions expressed are those of the authors and do not necessarily represent official positions or policy of the OFR or the U.S. Department of the Treasury. simpson.zhang@ofr.treasury.gov
misguided investments by financial firms, collapses suddenly and drastically. Why is this pattern so persistent? And why don’t financial firms more actively rein in potentially harmful investments by their managers?¹

This paper explains why compensation contracts with the “wrong” incentives are so prevalent during bubbles.² We consider a model in which two principals compete for the managerial talents of an agent who must choose between a safe and a risky action. The principals consider the safe action better than the risky action, but the agent is overly-optimistic and overvalues the risky action.³ We derive equilibrium contracts and show how intense labor market competition can shift the bargaining power to the agent and cause contracts to provide high-powered incentives that promote risk-taking. We identify three specific factors that can intensify competition and result in riskier contracts: higher job matching efficiency, larger project investments, and greater debt levels.

Prior empirical work has shown that competition for managers is enormous during major financial bubbles. For instance, Philippon & Reshef (2012) shows that the excess wage (wage controlling for education) received by the financial industry overall relative to other industries varies considerably over time, and in particular it skyrocketed prior to both the Great Depression and the 2008 financial crisis. The authors attribute this increase in excess wage to greater rents accrued by this sector, which provides evidence for the increase in competition for financial managers during these two time periods. The authors also show that excess executive compensation for financial executives underwent a large increase prior to the 2008 financial crisis.

A key part of our model is the incorporation of overly-optimistic manager beliefs. Overop-
timism is a common psychological phenomenon, and its prevalence during financial bubbles has been extensively documented (as famously stated through the phrase “irrational exuberance”). For instance, before the financial crisis financial managers commonly used flawed pricing models, such as Gaussian copula models, which did not accurately model correlations among asset prices. Many investors thus believed products such as CDOs were much safer than they were in reality. These models remained popular in spite of repeated warnings from technical experts that they were unsuitable for use in risk management\(^4\). As evidence of optimistic beliefs by these investors, Cheng et al. (2014) shows that securitization investors increased their personal housing exposures pre-crisis, and their housing portfolios performed much worse than those of control groups.

Subprime lending before the crisis also dramatically highlights how financial managers can make damaging investments due to their overoptimism, while financial firms do not prevent this behavior. Prior to the crisis, the subprime mortgage market featured substantial increases in lending to borrowers across the credit spectrum\(^5\), including those with very poor credit and low probability of repayment. Particularly egregious examples included NINJA loans, which were made to individuals with “No Income, No Job, No Assets”\(^6\). These loans were justified on the basis of optimistic beliefs about sustained growth in housing prices\(^7\). When housing prices tumbled instead, these loans collapsed.

Our paper helps explain these events by presenting a non-preference-based reason for why compensation contracts promote excessive risk-taking: market competition. When labor-market competition is intense, such as during financial bubbles, our model predicts that compensation contracts will incentivize more risk-taking. In particular, we show that agent overoptimism alone is not sufficient and that speculative motives by the principal are not necessary (as opposed to what’s guessed in Baron & Xiong (2017)), for such contracts to

\(^5\)See Adelino et al. (2016) for an empirical analysis.  
\(^6\)See for instance [http://www.telegraph.co.uk/finance/economics/2785403/Ninja-loans-explode-on-sub-prime-frontline.html](http://www.telegraph.co.uk/finance/economics/2785403/Ninja-loans-explode-on-sub-prime-frontline.html)  
\(^7\)Foote et al. (2012) provides evidence for this behavior.
emerge.

**Outline of the Paper** Section 2 describes the model. A risk-neutral principal wants to hire a risk-neutral agent to make an investment, which can be either safe or risky, by offering a contract with bilateral limited liability\(^8\). The returns from this investment are stochastic and depend on the action taken. The principal thinks the safe action has a higher expected payoff but the agent, who is overly-optimistic, thinks the risky action has a higher expected payoff. In the base model, actions are assumed to be contractable. The agent has a reservation utility, which we first assume is exogenous and we then endogenize by embedding the baseline model into a competing principals framework. To model competition, we consider a class of matching processes parameterized by a matching efficiency, which represents the intensity of competition in the labor market.

In Section 3, we solve the model. Why would the principal ever want to implement the risky action? The trade-off is that, by implementing the risky action, the principal can write a contract that pays the agent less than what the agent expects. Exploiting the agent’s heterogeneous beliefs is profitable for the principal when the agent’s reservation utility is high. In the competing principals framework, we show that higher matching efficiencies lead to higher equilibrium utility levels for the agent. For this reason, the risky action will be implemented in equilibrium if the matching efficiency is high.

In Section 4 we discuss two extensions of the model. First, we show that increases in the fund scale can cause equilibrium actions to become risky. A larger fund scale means a successful hire is more profitable for the principals, which intensifies the competition for managerial talents. Environments in which investments are large, such as financial bubbles, will thus feature contracts that promote excessive risk-taking. Second, we show that loosening the limited liability constraints and letting the principal take debts allows the principal to more profitably exploit the agent’s optimistic beliefs. These higher profits can also increase competition and lead to more risk-taking in equilibrium. Greater firm leverage levels are thus another reason contracts may incentivize risky actions.

\(^8\)Our use of bilateral limited liability is similar to Innes (1990).
In Section 5, we discuss the implications of hidden actions for our results. We show that when a principal cannot contract on the agent’s action, profitability decreases and the safe action may no longer be implementable at high utility values. However, hidden actions can also reduce the level of competition between principals. The equilibrium effect of hidden actions can thus be beneficial, and safe actions may be sustainable for a wider range of matching efficiencies than before. Our result indicates that corporate governance policies that mandate greater manager oversight need to carefully consider their equilibrium implications.

Finally, section 6 concludes the paper. Longer proofs are contained in the appendix.

The Literature Our paper builds on and bridges together several disparate strands of the literature. First, we contribute to the literature on bank CEO incentives.\footnote{For a recent survey of this vast literature, see Edmans & Gabaix (2016).} There is currently a great debate over the role compensation contracts played in the financial crisis. One side argues that compensation contracts created poorly aligned incentives\footnote{For example, the contract is chosen to maximize CEO rents instead of shareholder value.} that encouraged excessive risk-taking and thus contributed to the crisis (Bebchuk & Spamann (2010), Bebchuk & Fried (2009)). The other side (Fahlenbrach & Stulz (2011)) asserts that contracts did provide proper incentives as many CEOs had a large amount of their equity at stake. Under the second view, behavioral reasons provide a possible explanation for the bad investments. Our work connects these two views: heterogeneous beliefs and labor market competition work jointly to cause firms to offer compensation contracts that incentivize suboptimal decisions.

Our paper also contributes to the study of labor market competition in agency models. Bénabou & Tirole (2016) features a model in which the agent exerts effort on a non-contractable and a contractable task, with the output from the latter being talent-dependent. Competition causes firms to offer higher incentives and bonuses to screen for talent, thereby driving the incentives for the contractable task from downward distorted to upward distorted relative to the optimum. In our paper, heterogeneous beliefs instead of screening motives
are the main consideration in the principal’s contracting problem. Competition results in riskier actions that actually reduce manager compensation \textit{ex post} due to the manager’s overly-optimistic beliefs about the contract’s payoff. This result is in line with the empirical findings of Otto (2014), which shows that optimistic managers receive less total compensation and fewer bonuses than peers. Another relevant theoretical paper is Acharya et al. (2016), in which the ability of the agent is gradually revealed if the agent stays in the same firm long enough. Competition is captured by labor mobility, and mobility offers the low talent agent a way to delay being discovered as low by moving to other firms. Most closely related in spirit to our paper is Thanassoulis (2013). In his model the optimal contract mitigates the impatient agent’s short-termism induced moral hazard by delaying bonuses. With competition, firms offer more short-term bonuses, increasing the agents’ short-termism. Our paper considers both a different source of disagreement, caused by heterogeneous beliefs, and a different model of competition that incorporates matching frictions. We show that higher matching efficiencies lead to more risk-taking. Crucially, in our model with competition, moral hazard can actually limit the amount of risk-taking in equilibrium.

Our paper is also related to optimal contract design with heterogeneous beliefs.\textsuperscript{11} Gervais et al. (2011) analyzes contract design with an agent who is both risk-averse and overconfident.\textsuperscript{12} Risk-aversion may prevent the agent from taking the amount of risk the principal desires. However, the optimal contract takes advantage of the agent’s overconfidence and implements the principal-preferred action. Bolton et al. (2006) find that when the principal thinks the market is speculative, the optimal incentive contract directs the manager to spend more effort on castle-in-the-air projects because of the arbitrage opportunity. Palomino & Sadrieh (2011) considers a model with information acquisition and analyze how overcon-

\textsuperscript{11}Goel & Thakor (2008) shows that the CEO selection process tends to favor overconfident agents to become CEOs.

\textsuperscript{12}Overconfidence is typically modeled as the agent having a higher subjective signal precision than in reality regarding a risky investment. Overoptimism, on the other hand, is having a higher subjective probabilistic belief that good outcomes occur. In the appendix we show that our model can also be formulated in terms of overconfidence.
fidence shapes the optimal contract. Wang et al. (2013) considers a similar model with optimism instead of overconfidence. De la Rosa (2011) studies optimal contract design when the principal and the agent have heterogeneous beliefs about how effort affects project success. Sautmann et al. (2013) provides experimental evidence in a similar environment that principals take advantage of agents’ overconfidence.

Most of these contract design papers assume that the principal prefers riskier actions than the agent. Under such conditions the prevalence of risky actions in equilibrium may not seem surprising. However, we use the competition for managerial talents to explain the prevalence of contracts that promote risk-taking even in contexts where the principal doesn’t want the agent to take the risky action in the first place. Our model thus shows that even under the most conservative set of assumptions, the forces of competition can still cause excessive risk-taking.

Finally, our paper has implications for the leverage cycle. Geanakoplos (2010) argues that the leverage cycle is driven by optimistic or pessimistic beliefs due to the occurrence of good or bad news. Our model suggests that competition for managers, which is high in boom times and low in bad times, amplifies the leverage cycle.

2 The Model

Players Risk-neutral principal, risk-neutral agent. Principal (she) hires an agent (he) to manage an investment.

Actions Hired agent takes action $a \in \{s, r\}$, where $s$ is the safe action and $r$ is the risky venture.

Beliefs The principal thinks that investment payoffs have densities $f(x|a)$, $a \in \{s, r\}$. The agent agrees that the safe investment generates payoffs according to $f(x|s)$, but believes that the risky venture generates payoffs according to density $f_A(x|r)$ instead. Densities are continuously differentiable and everywhere positive. Most importantly, we assume that

$E[x|r] < E[x|s] < E_A[x|r]$. 

We interpret the safe action as representing the common and well-known action, so the two parties have no disagreement over its payoff. The risky action on the other hand represents the newer or more complex investment. Although possible, we do not require the payoff from the risky action to have a higher variance than that of the safe action. Instead, “riskiness” in our model is broadly interpreted as meaning easier to misunderstand due to its novelty or complexity. For instance prior to the financial crisis, although many Collateralized Debt Obligations (CDOs) were highly rated, their inherent complexity may them difficult to value correctly, and their use could thus be considered risky.

**Contract** A contract is a pair \((w, a)\) where \(w : [0, 1] \to [0, 1]\) is a measurable function with bilateral limited liability: \(0 \leq w(x) \leq x\), and \(a \in \{r, s\}\). We call a contract a safe contract if \(a = s\) and a risky contract if \(a = r\).

**Utilities** Suppose the contract \((w, a)\) is offered, accepted, and action \(a\) is taken. The principal’s expected utility is

\[\Pi_a = E[x - w(x)|a]\]

The agent’s expected utility for accepting the contract is

\[u = E_A[w(x)|a],\]

where \(E_A[w(x)|s] = E[w(x)|s]\).

The optimal profit for implementing action \(a\) when the agent’s reservation utility is \(u\) is

\[\Pi_a(u) = \max_{w(\cdot)} E[x - w(x)|a]\]

such that \(w(\cdot)\) is measurable, \(0 \leq w(x) \leq x\) and that \(E_A[w(x)|a] \geq u\).

**Competing Principals and the Matching Function** The above baseline model can be embedded into a competing principals framework as follows:

Two principals, \(i = 1, 2\), compete for one agent by offering contracts \((w_1, a_1), (w_2, a_2)\). Let \(u_1 = E_A[w_1|a_1], u_2 = E_A[w_2|a_2]\) be the utility that the agent receives from each contract.

\(^{13}\)Bilateral limited liability guarantees that a solution exists when both sides are risk-neutral. We relax this condition in Section 5.
The probability that principal $i$ is matched with an agent depends on a matching function $p : [0, \infty)^3 \to (0, 1]$, with matching efficiency parameter $\beta$, such that $p \in C^2$ and it satisfies:

**Assumption 1.**

1. $p(u_1, u_2; \beta) + p(u_2, u_1; \beta) = 1$

2. $p_1 \geq 0, p_3 > 0$ when $u_1 > u_2$.

3. $\partial \ln p/\partial u_1$ is decreasing in $u_1$, $\partial \ln p/\partial u_2$ is increasing in $u_1$. Furthermore, $p_1/p = 0$ when $\beta = 0$, is increasing in $\beta$ when $u_1 \leq u_2$, and along $u_1 = u_2 = u$, $p_1/p$ is non-increasing in $u$ and increases from zero to infinity when $\beta$ goes from zero to infinity.

Note that 1. and 2. imply that $p_2 \leq 0, p_3 < 0$ when $u_1 < u_2$, and $p_3 = 0$ when $u_1 = u_2$.

Item 2 says that the matching probability is increasing in own offering and increasing in the matching efficiency when own offering is higher.

Item 3 says that the marginal effect of own offering on matching probability does not increase too much when own offering increases, does not decrease too much when other’s offering increases, and the own offering elasticity, when evaluated at $u_1 = u_2 = u$, does not increase too fast in $u$, and spans the entire positive reals when the matching efficiency parameter $\beta$ varies across the positive reals.

Our assumptions on the matching problem represent a natural set of restrictions. They mainly entail that the matching function shows decreasing returns to scale in offering higher levels of utility.

We give an example of a matching process that satisfies our assumptions.

**Example 2.1 (Normal Noise on Wage Offered).** Let $\{\epsilon_i\}$ be i.i.d. random variables distributed as $N(0, \frac{1}{\beta^2})$. Suppose principal $i$ offers a contract with indirect utility $u_i$ to some type of agent. Then the contract generates a signal $s_i = u_i + \epsilon_i$. Agents are matched to the principal with the highest realized signal.

Hence, given $u_1, u_2$,

$$p(u_1, u_2; \beta) = 1 - \Phi(\sqrt{\beta}(u_2 - u_1))$$
and

\[
\begin{align*}
    p_1(u_1, u_2; \beta) &= \frac{\sqrt{\beta} \phi(\sqrt{\beta}(u_2 - u_1))}{1 - \Phi(\sqrt{\beta}(u_2 - u_1))}, \\
    p_2(u_1, u_2; \beta) &= -\frac{\sqrt{\beta} \phi(\sqrt{\beta}(u_2 - u_1))}{1 - \Phi(\sqrt{\beta}(u_2 - u_1))}, \\
    p_3(u_1, u_2; \beta) &= \frac{(u_2 - u_1) \phi(\sqrt{\beta}(u_2 - u_1))}{2 \sqrt{\beta}(1 - \Phi(\sqrt{\beta}(u_2 - u_1)))}.
\end{align*}
\]

which indeed satisfies our assumptions on the matching function.

**Equilibrium** Our equilibrium notion is pure-strategy Nash equilibrium:

**Definition 2.1.** A pure strategy Nash equilibrium is a pair of contracts \((w_1, a_1), (w_2, a_2)\) such that \(E_A[w_i|a_i] = u_i\) and \((w_1, a_1)\) solves

\[
\max_{(w', a')} p(u', u_2; \beta) \Pi_{a'}(u')
\]

s.t.

\[E_A[w'|a'] = u'\]

and likewise for \((w_2, a_2)\).

We will focus on symmetric pure-strategy Nash equilibrium in the analysis (equilibrium in which principals implement the same action \(a\) with the same expected wage \(E_A[w|a]\)), and we show in Appendix B that asymmetric pure-strategy Nash equilibria do not exist. A symmetric pure-strategy Nash equilibrium is called a safe equilibrium if \(a = s\), and a risky equilibrium if \(a = r\).

### 3 Contracts and Competition

In this section we characterize the optimal contract when there is a single principal, and we then extend the model to one with competing principals and characterize the Nash equilibrium.
3.1 Optimal Contracts with Single Principal

When there is a single principal, we assume that the agent accepts a contract if and only if the contract offers equal or higher utility than his reservation utility.

It is easily seen that
\[ \Pi_s(u) = E[x|s] - u \]
and any wage scheme such that \( E[w(x)|s] = u \) is optimal. It is also easily seen that \( \Pi_s(E[x|s]) = 0 < \Pi_r(E[x|s]) \) because \( E_A[x|r] > E[x|s] \) implies that when implementing the risky action, the principal need not give the agent the entire project to satisfy his reservation of \( u = E[x|s] \).

The main driving force of our model is the following comparison of optimal profits:

**Proposition 3.1.** \( \Pi_r(u) \) is decreasing and weakly concave. In particular, there exists \( u^* < E[x|s] \) such that \( \Pi_s(u) > \Pi_r(u) \) if \( u < u^* \) and \( \Pi_s(u) < \Pi_r(u) \) if \( u > u^* \).

This result has a strong intuition: while the principal perceives the safe action to be the superior one, if it is relatively too costly to hire the agent to do so, the principal would rather implement the risky action to take advantage of the agent’s “wrong” belief. The more competitive the labor market is, the higher the wages, and as a result the more risky actions will be taken due to the need to match the competition.

To see why \( \Pi_r \) is concave, let \( w_i \) be the optimal wage implementing \( a = r \) when the agent’s reservation utility is \( u_i, i = 1, 2 \). Then for any \( u = \rho u_1 + (1 - \rho)u_2, \rho \in (0, 1) \), by linearity of the agent’s payoff function, the wage scheme \( \rho w_1 + (1 - \rho)w_2 \) gives the agent utility \( u \). However, it may not be principal-optimal. Therefore \( \Pi_r(u) \geq \rho \Pi_r(u_1) + (1 - \rho)\Pi_r(u_2) \).\(^{14}\) Since \( \Pi_r(0) < \Pi_s(0) \) and \( \Pi_r(E[x|s]) > \Pi_s(E[x|s]) = 0 \), \( \Pi_r \) crosses \( \Pi_s \) exactly once, from below. **Figure 1** presents an example that highlights the optimal profit obtainable under each type of contract for different reservation utilities.\(^{15}\)

\(^{14}\)In fact, the optimal contract pays the agent \( x \) whenever \( \lambda f_A(x|r) > f(x|r) \), where \( \lambda \) is the Lagrange multiplier. See Appendix A.1. for a full characterization.

\(^{15}\)The beliefs used in the example are: \( f(x|s) = 1, f_A(x|r) = 2x, f(x|r) = .5x^{-5} \). They satisfy MLRP dominance and therefore also satisfy expected value dominance.
3.2 Competition

In this subsection we characterize the pure strategy Nash equilibria in a competitive labor market. The competition among principals is parameterized by an exogenous matching efficiency parameter $\beta$. We show that as $\beta$ goes up principals compete more fiercely in wages by giving the agent better wage contracts. Combined with Proposition 3.1, the equilibrium contract will be risky once the matching efficiency is sufficiently high. We focus on symmetric equilibria and show in Appendix B that asymmetric equilibria do not exist.

For each $a \in \{s, r\}$, the expression $p(u_i, u_{-i}; \beta)\Pi_a(u_i)$ defines an auxiliary normal form game. Observe that the utility offered to the agent by any equilibrium $(w, a)$ in the competing principals game, $E_{A}[w(x)|a]$, must be a Nash equilibrium of the corresponding auxiliary game.

In the auxiliary game, the best-response of player $i$ to any $u_{-i}$ satisfies the FOC

$$p_1(u_i, u_{-i}; \beta)\Pi_a(u_i) + p(u_i, u_{-i}; \beta)\Pi_a'(u_i) = 0.$$ 

Since $p$ is assumed to be log-concave in $u_i$, the symmetric interior Nash equilibrium is therefore
characterized by
\[ \frac{p_1(u, u; \beta)}{p(u, u; \beta)} = \frac{-\Pi'_s(u)}{\Pi_a(u)}. \] (1)

The LHS of (1) is a non-increasing function of \( u \) and the RHS of (1) is an increasing function of \( u \). Denote the Nash equilibrium by \( u_a(\beta) \). The assumptions on \( p \) say that the LHS shifts to the right as \( \beta \) increases, leading to a higher equilibrium \( u_a(\beta) \) (see Figure 2).

\[ \begin{align*}
\pi_1(u, u; \beta) &\quad \pi_1(u, u; \beta') \\
\pi(u, u; \beta) &\quad \pi(u, u; \beta') \quad \Pi_s(u) \\
\beta \uparrow \beta' &\quad 1 \quad E[x|s]
\end{align*} \]

Figure 2: Equilibrium Utility of Safe Contracts

Intuitively, as \( \beta \) increases, \( u_s(\beta) \) will surpass \( u^* \), in which case offering a risky contract with the same utility is more profitable for the principal. In fact, as we show later, the safe equilibrium cannot be sustained even before \( u_s(\beta) \) reaches \( u^* \).

To capture the profitability of such deviations for any \( \beta \), define four value functions as...
follows.

\[ E\Pi_s(\beta) = p(u_s(\beta), u_s(\beta); \beta)\Pi_s(u_s(\beta)) \]

\[ E\Pi'_s(\beta) = \max_{u \in [0, E_A[x|r]]} p(u, u_s(\beta); \beta)\Pi_r(u) \]

\[ E\Pi_r(\beta) = p(u_r(\beta), u_r(\beta); \beta)\Pi_r(u_r(\beta)) \]

\[ E\Pi'_r(\beta) = \max_{u \in [0, E[x|s]]} p(u; u_r(\beta); \beta)\Pi_s(u) \]

Also, define \( u^*_s(\beta) \) to be the maximizer of \( E\Pi'_s(\beta) \) and \( u^*_r(\beta) \) to be the maximizer of \( E\Pi'_r(\beta) \).

Therefore, for each \( \beta \), \( E\Pi_s(\beta) \) is the on-path profit if everyone offers a safe contract with utility \( u_s(\beta) \) when the matching efficiency parameter is \( \beta \), and \( E\Pi'_s(\beta) \) is the greatest profit from a deviation to a risky contract when the opponent is still offering a safe contract with utility \( u_s(\beta) \). Likewise for \( E\Pi_r, E\Pi'_s \).

We characterize the Nash equilibria of the competing-principals game via these value functions in the following lemma.

**Lemma 3.1.** \((w, a)\) is a symmetric Nash equilibrium of the full-information competing-principals game with matching efficiency parameter \( \beta \) if and only if

\[ E_A[w(x)|a] = u_a(\beta) \]

\[ E\Pi_a(\beta) \geq E\Pi'_a(\beta) \]

where \( a, a' \in \{s, r\}, a \neq a' \) and \( u_a(\beta) \) is the equilibrium in the corresponding auxiliary game.

**Proof.** Follows directly from definitions. \(\square\)

The next lemma shows that there are profitable risky deviations from offering the safe contract if and only if \( \beta \) is higher than a threshold \( \beta_s \). Similarly, there are profitable safe action deviations from a risky contract if and only if the competition intensity is below a threshold \( \beta_r \).

**Lemma 3.2.**

\( E\Pi_s(\beta) \) crosses \( E\Pi'_s(\beta) \) exactly once, from above at some \( \beta_s \in (0, \infty) \).
\( \Pi_r(\beta) \) crosses \( \Pi_s(\beta) \) exactly once, from below at some \( \beta_r \in (0, \infty) \).

Moreover, \( \beta_r < \beta_s \).

The intuition is simple. When \( \beta \) increases, a deviation from the safe equilibrium to the risky action becomes more attractive because larger wages are more effective at leading to a successful hire, and \( \Pi_r(u) \) eventually becomes higher than \( \Pi_s(u) \). Likewise, when \( \beta \) decreases, a deviation from the risky equilibrium to the safe action becomes more attractive due to the same logic working backwards.\(^{16}\)

We can now state our first main result:

**Theorem 1.** Under full information, there exist \( 0 < \beta_r < \beta_s \) such that

1. A safe equilibrium exists if and only if \( \beta \leq \beta_s \).

2. A risky equilibrium exists if and only if \( \beta \geq \beta_r \).

**Proof.** Follows from Lemma 3.1 and Lemma 3.2. \( \square \)

That is, safe action equilibria exist only when competition is weak, and risky action equilibria exist only when competition is intense. There is also an intermediate region of competition where both types of equilibria are possible.

As one implication of our result, policies such as CEO compensation disclosure or wage transparency will potentially lead to more intense competition among the principals. These could lead to, aside from higher CEO pay, more risk-taking.

\(^{16}\)We use an envelope theorem argument to show the desired single-crossing. However, the differentiability of \( u_s(\cdot), u_r(\cdot) \) is in general not guaranteed, and this leads to some technicalities that we deal with in the appendix.
4 Other Factors That Lead to Risk-Taking

4.1 Fund Scale and Excessive Risk-Taking

What is the relationship between the size of a fund and its performance? Chen et al. (2004) finds that the scale of a mutual fund erodes its performance due to liquidity and organizational diseconomies. In addition to liquidity, Pollet & Wilson (2008) finds funds respond to growth by increasing the number of shares already held rather than diversifying the portfolio, which leads to diminishing returns of scale.

Our model shows a similar result, via a different channel: competition. As the fund scales up, the competition for managerial talents intensifies because the payoffs of a successful hire to a principal increase. If the equilibrium wage is driven up disproportionately more as fund scale increases, the market equilibrium will switch from safe to risky contracts.

Specifically, suppose that the fund size is some \( \alpha \geq 1 \) and that there is no diversification option but to increase the number of shares for the same investment. The returns of the investment are now \( \alpha x \), where \( x \) has the same densities \( f(x|s), f(x|r), f_A(x|r) \) as before. Let \( \Pi_s(u, \alpha), \Pi_r(u, \alpha) \) denote the corresponding optimal principal payoffs when implementing \( a = s \) or \( a = r \) respectively. They are characterized exactly the same as in Appendix A.1. except that there is now a scalar \( \alpha \).

Let \( u_s(\beta, \alpha) \) and \( u_r(\beta, \alpha) \) be the symmetric equilibria in the corresponding auxiliary game \( p(u_1, u_2; \beta)\Pi_s(u_1, \alpha) \) and \( p(u_1, u_2; \beta)\Pi_r(u_1, \alpha) \). It follows directly from the FOCs that \( u_s \) and \( u_r \) are increasing in \( \alpha \) whenever \( u_s, u_r > 0 \).

How much the competitive pay (equilibrium indirect utility) increases with respect to \( \alpha \) depends on the matching function. In this subsection we assume, in addition to Assumption 1, that the matching probability depends only on the difference of the indirect utilities offered.

Assumption 2. There exists a function \( q : \mathbb{R} \times \mathbb{R}_+ \rightarrow (0, 1] \) such that \( p(u_1, u_2; \beta) = q(u_1 - u_2; \beta) \) for all \( u_1, u_2 \) and \( \beta \).
Remark 4.1. Note that Assumption 2 implies that
\[
\frac{p_1(u, u; \beta)}{p(u, u; \beta)}
\]
is constant in \(u\) and that
\[
\frac{p_1(u_1, u_2; \beta)}{p(u_1, u_2; \beta)} = -\frac{p_2(u_1, u_2; \beta)}{p(u_1, u_2; \beta)}
\]
for all \(u_1, u_2, \beta\).

In particular, the normal noise case of Example 2.1 satisfies this assumption. One interpretation of this assumption is that the agent values additional utility the same at all levels of utility, and thus only considers the difference in utility when choosing a firm. For instance, in the normal noise example the agent’s utility may be the sum of the utility of the contract plus an idiosyncratic firm-specific utility \(\varepsilon_i\). The agent goes to whichever firm offers it the highest overall utility, but the firms do not observe the values of \(\varepsilon_i\) when making their offers.

The equilibrium indirect utility \(u\) is determined by the FOC
\[
\frac{p_1(u, u; \beta)}{p(u, u; \beta)} = -\frac{\Pi'(u, \alpha)}{\Pi(u, \alpha)}
\]
The left-hand side is downward sloping and the right-hand side is upward sloping. An increase in \(\alpha\) shifts the right-hand side curve to the right and therefore increases \(u\). Assumption 2 implies that the left-hand side is flat, maximizing the impact of the increase of \(\alpha\) on the equilibrium indirect utility.

Under Assumption 2, we can generalize Theorem 1 as follows.

**Theorem 2.** Under full information, for every \(\alpha \geq 1\) there exists \(0 < \beta_r(\alpha) < \beta_s(\alpha)\) such that

1. A safe symmetric equilibrium exists if and only if \(\beta \leq \beta_s(\alpha)\)

2. A risky symmetric equilibrium exists if and only if \(\beta \geq \beta_r(\alpha)\).

Furthermore, \(\beta_s(\cdot), \beta_r(\cdot)\) are decreasing, and \(\lim_{\alpha \to \infty} \beta_s(\alpha) = 0\).

In particular, for any \(\beta > 0\), as long as the scale is large enough, the risky equilibrium is the only equilibrium.
4.2 Limited Liability Constraints and Excessive Risk Taking

What is the implication for equilibrium contracts if we allow either the principals or the agent to take debts?

First, if we allow the agent to take debts, then implementing the risky action could be optimal even for low values of reservation utility, because the principal can take advantage of the heterogeneous beliefs by making a bet such that agent pays the principal when the return of the project is low.

Formally, we relax the limited liability constraint to

\[-L \leq w(x) \leq x\]

where \( L \) is the parameter of the amount of debt the agent can take. Then we have the following result.

**Theorem 3.** Consider the problem

\[
\max_{w(x)} E[x - w(x)|r]
\]

s.t.

\[-L \leq w(x) \leq x\]

\[E_A[w(x)|r] = u\]

Let \( \Pi_r(u, L) \) denote its value function. Then there exists an \( L^* \) such that for all \( L > L^* \),

\[\Pi_r(u, L) > E[x|s] - u\]

for all \( u \). Furthermore, the optimal wage scheme satisfies IC.

A direct implication is that when \( L \) is large enough, there will be no safe equilibrium in the competing principals game, regardless of \( \beta \).

Next, suppose the principal can take debts \( L \). That is,

\[0 \leq w(x) \leq x + L\]
A higher $L$ increases the principal’s payoff from implementing the risky action, for each level of reservation utility. However, it is no longer the case that $\Pi_r(L, u) > \Pi_s(u)$ for all $u$ for sufficiently high $L$: A higher $L$ means that the principal can promise the agent more payoff. But if $u$ is low, there is no need to promise the agent a high payoff. The principal’s payoff from implementing the safe action is unchanged: $\Pi_s(u, L) = E[x|s] - u$.

We have the following generalization of Theorem 1:

**Theorem 4.** Suppose the limited liability constraint is $0 \leq w(x) \leq x + L$. There exists $0 < \beta_r(L) < \beta_s(L)$ such that

1. A safe equilibrium exists if and only if $\beta \leq \beta_s(L)$
2. A risky equilibrium exists if and only if $\beta \geq \beta_r(L)$.

Furthermore, both $\beta_s(\cdot)$ and $\beta_r(\cdot)$ are decreasing in $L$.

5 **Discussion: Hidden Action**

Up to now, we have assumed that the agent’s action can be monitored by the principal. However, in some circumstances such monitoring may be too costly or even impossible to implement. Thus, in this section we discuss the potential implications of hidden actions on our results.

We show surprisingly that hidden actions can actually improve social welfare. With only a single principal, hidden actions will reduce profits and efficiency because it restricts the range of safe action contracts that can be offered. However, with competition having hidden actions can be beneficial because it limits the intensity of competition among the principals. Since hiring the agent is no longer as profitable, principals may compete less for the agent and offer lower utilities in equilibrium. We show that in certain situations, it is even possible for safe action equilibria to exist for any value of the matching efficiency parameter $\beta$. 

19
5.1 Optimal Profit

We first characterize the optimal profit of a single principal when the reservation utility $u$ of the agent is exogenously given.

The principal’s contract design problem is

$$
\max_{(w(\cdot), a)} \int_0^1 (x - w(x)) f(x|a)dx
$$

subject to

$$
E_A[w(x)|a] \geq u \quad \text{(IR)}
$$
$$
E_A[w(x)|a] \geq E_A[w(x)|a'] \quad \text{(IC)}
$$
$$
0 \leq w(x) \leq x \quad \text{(LL)}
$$

Define $\tilde{\Pi}_a(u)$ as the value function of the principal when she implements action $a \in \{s, r\}$. A wage scheme $w$ implements action $a$ if $w$ satisfies IR, IC and LL.

First, we claim that due to the IC constraint, the safe action is now implementable when $u \in [0, \tilde{u}]$ for some $\tilde{u} < E[x|s]$. The profit over this range will still be the same as before.

To see this, note that when the reservation utility is $u = E[x|s]$, the principal has to give a wage scheme such that $w(x) = x$ a.s. But the assumption that $E_A[x|r] > E[x|s]$ implies that the agent will pick the risky action. Furthermore, whenever $w, w'$ implement $s$ with reservation utilities $u, u'$ respectively, the convex combinations of $w$ and $w'$ implements $s$ with reservation utilities inbetween $u$ and $u'$.

Second, the above convex combination argument shows that $\tilde{\Pi}_r(u)$ is concave and that the risky action is still implementable for any reservation utility $u \in [0, E_A[x|r]]$. However, it may be that $\tilde{\Pi}_r(u) < \Pi_r(u)$ for some $u$ due to the additional IC constraint.

There are two possible cases. First, $\tilde{\Pi}_s(\tilde{u}) > \tilde{\Pi}_r(u)$. Second, there exists $u^* < \tilde{u}$ such that $\tilde{\Pi}_s(u^*) = \tilde{\Pi}_r(u^*)$. Figure 3 shows the two cases.

Remark 5.1. Whether $\tilde{u} < u^*$ depends on the configuration of the beliefs. Recall that $u^*$ is the cutoff of agent’s reservation utility above which the optimal contract is risky. Therefore,
if action \( r \) is perceived to be very inferior by the principal, then the principal will not want to design a risky contract unless the reservation utilities are really high, so \( u^* > \tilde{u} \). On the other hand, if the risky action is perceived to be just slightly inferior by the principal, then the principal would like to take advantage of the agent’s “wrong” belief even if his reservation utility is low, so \( u^* < \tilde{u} \).\(^{17}\)

![Figure 3: Optimal Profits With Hidden Action](image)

5.2 Competition

The fact that principals who want to implement the safe action cannot offer a utility higher than \( \tilde{u} \) has non-trivial implications for the set of equilibria under \( \beta \).

First, suppose the principals offer a safe contract and \( \beta \) is high enough so that the utility to the agent is \( \tilde{u} \). If a principal wants to deviate and offer the agent a higher utility, she has no

\(^{17}\)As a numerical example, let \( f(x|s) = 2x \) and \( f_A(x|r) = 3x^2 \). This implies \( \tilde{u} = 0.468 \). If \( f(x|r) = 1 \), then \( u^* = 0.26 < \tilde{u} \). If \( f(x|r) = 10(1-x)^9 \) then \( u^* = 0.58 > \tilde{u} \).
choice but to offer a risky contract. This entails a discrete loss of profit when $\Pi_s(\tilde{u}) > \Pi_r(\tilde{u})$. Therefore, if the increase in the probability of hiring the agent does not compensate this loss, both principals would prefer to offer the safe contract.

This may be true even when $\beta$ becomes arbitrarily large. For example, suppose $0.5\Pi_s(\tilde{u}) \geq \Pi_r(\tilde{u})$. Note that $0.5\Pi_s(\tilde{u})$ is the expected profit of offering the safe contract and $\Pi_r(\tilde{u})$ is the upper bound of the profit of offering the risky contract when $\beta$ is large because the matching probability of any risky action deviation is bounded above by 1. Under such a set of parameters, safe action equilibria can now be supported for all levels of $\beta$.

This example shows how hidden actions can increase efficiency with competition. Intuitively, hidden actions restrict the amount of competition between principals and lower the utilities that are offered in equilibrium. This in turn increases the range of safe action contracts that can be supported.

We note that the set of risky action equilibrium can also be affected by hidden actions. Suppose the principals both offer risky contracts. Deviations to safe contracts must offer the agent no more than $\tilde{u}$ utility, making such deviations more restrictive and less attractive. However, the IC constraint also implies that $\Pi_r(u) \leq \Pi_r(\tilde{u})$, and this potentially makes offering risky contracts less attractive as well. The overall effect on the set of $\beta$ that supports the risky equilibrium is therefore ambiguous.

Another technical problem is that $\tilde{\Pi}_r(u)$ may be non-differentiable at $u$’s where the IC constraint changes from binding to non-binding. This further complicates the equilibrium analysis. One condition that eliminates such complications is when $f(x|s)$ MLRP dominates $f(x|r)$ and $f_A(x|r)$ MLRP dominates $f(x|s)$. Under this assumption, it can be shown that the optimal risky contract will always be of the form $w(x) = x$ whenever $x \geq \bar{x}$ and $w(x) = 0$ otherwise, and the IC constraint will always be strict. This implies that $\tilde{\Pi}_r(u) = \Pi_r(u)$ for all $u$, and that the set of $\beta$ that supports the risky equilibrium becomes weakly larger.
6 Conclusion

This paper studies the effects of labor market competition on risk-taking in a heterogeneous belief model. Due to limited liability, the principal cannot use side bets to take advantage of the heterogeneous belief. Hence, the contract offered by the principal is determined by the pressure to offer the agent a greater share of the project. High-powered incentive contracts are used when the competition for managerial talent is intense, as in a financial bubble.

We assumed risk-neutral utilities and heterogeneous beliefs for ease of tractability. A model with risk-averse principals and agents in general does not have an analytical solution and the relevant cutoffs on the reservation utility will be difficult to pin down. However, the qualitative results in our paper do not seem to depend on risk-neutrality.

Our paper highlights the potential reasons why optimistic beliefs and bad investments seem to prevail so frequently in financial bubbles. The higher levels of labor market competition in a bubble naturally lead to greater levels of manager optimism and risk-taking. Regulations which seek to improve corporate governance and monitoring of manager decisions must be careful of inadvertently increasing a firm’s competitive drive and risk-taking motivations even further.
A Proofs of Main Results

We first establish Proposition 3.1.

Define

$$\mathcal{W} = \{ w : [0, 1] \to \mathbb{R} : w \text{ is Lebesgue measurable and } 0 \leq w(x) \leq x, \forall x \in [0, 1] \}. $$

Equip $\mathcal{W}$ with the $L^1$ norm so that $\mathcal{W}$ is complete and totally bounded, hence compact.\(^{18}\)

The optimal profit for a risky contract is given by

$$\Pi_r(u) = \max_{w \in \mathcal{W}} E[x|r] - E[w(x)|r] \tag{3}$$

s.t. $E_A[w(x)|r] \geq u$

The optimal wage exists because we are maximizing a linear function over a compact set. Note also that $\Pi_r(E[x|s]) > 0$ because $E_A[x|r] > E[x|s]$.

Proof of Proposition 3.1. To see that $\Pi_r$ is weakly concave on $[0, E_A[x|r]]$, let $0 \leq u_1 < u_2 \leq E_A[x|r]$. Let $\rho \in (0, 1)$. Let $w_1, w_2$ be the corresponding optimal wage. Then $\rho w_1 + (1 - \rho)w_2 \in \mathcal{W}$ and it satisfies the IR constraint when the reservation utility is $\rho u_1 + (1 - \rho)u_2$, but it may not be optimal. Therefore

$$\Pi_r(\rho u_1 + (1 - \rho)u_2) \geq E[x|r] - E[\rho w_1 + (1 - \rho)w_2|r] = \rho \Pi_r(u_1) + (1 - \rho)\Pi_r(u_2)$$

Now we can show that $\Pi_r$ and $\Pi_s$ cross exactly once. Since $\Pi_r$ is weakly concave, it is continuous on the interior of its domain, so $\Pi_r$ and $\Pi_s$ cross at some $u^*$. Assume to the contrary that there exist $u_1 < u_2$ such that $\Pi_r(u_1) = \Pi_s(u_1)$ and $\Pi_r(u_2) = \Pi_s(u_2)$, then $u_2 < E[x|s]$. Let $\rho$ be such that

$$u_2 = \rho u_1 + (1 - \rho)E[x|s].$$

\(^{18}\)A Cauchy sequence in $\mathcal{W} \in L^1([0, 1])$ has a convergent subsequence to some $w \in L^1([0, 1])$ because of $L^1$-completeness. The limit $w$ is s.t. $0 \leq w(x) \leq x$ a.e. and can be replaced with a $\tilde{w} \in \mathcal{W}$ s.t. $\tilde{w} = w$ a.e. Total boundedness is a consequence of the Kolmogorov–Riesz Compactness Theorem.
Then
\[ \rho \Pi_r(u_1) + (1 - \rho) \Pi_r(E[x|s]) > \rho \Pi_s(u_1) + (1 - \rho) \Pi_s(E[x|s]) \]
\[ = \Pi_s(u_2) = \Pi_r(u_2) \]
where the first equality follows from the linearity of \( \Pi_s \). This violates the fact that \( \Pi_r \) is weakly concave.

In what follows, we make preparations to prove Lemma 3.2. This involves two key steps:
1. Show that the \( u_r(\cdot) \), characterized by (1), has one-sided derivatives. \( u_s(\cdot) \) is easier since \( \Pi_s(u) \) is linear.) 2. Apply a one-sided version of the Envelope Theorem to \( E\Pi_a, E\Pi_a' \) and show single-crossing.

To show that \( u_r(\cdot) \) has one-sided derivatives, we need to show that the implicit function (1) that defines \( u_r(\cdot) \) has one-sided derivatives. That is, \( \Pi_r''(u+) \), \( \Pi_r''(u-) \) exist. Such a task can be achieved by a more refined characterization of the optimal wages that solve (3) in terms of the Lagrange multiplier of the constrained maximization problem, which in turn captures how the multiplier changes when the reservation utility changes.

Let
\[ m = \min_{x \in [0,1]} \frac{f(x|r)}{f_A(x|r)}, \quad M = \max_{x \in [0,1]} \frac{f(x|r)}{f_A(x|r)} \]

For each \( \lambda \in [m, M] \), define
\[ A_\lambda = \{ x \in [0,1] : \lambda f_A(x|r) > f(x|r) \} \]
\[ B_\lambda = \{ x \in [0,1] : \lambda f_A(x|r) < f(x|r) \} \]
\[ C_\lambda = \{ x \in [0,1] : \lambda f_A(x|r) = f(x|r) \} \]
\[ W_\lambda = \left\{ w \in W, w(x) = \begin{cases} x, & x \in A_\lambda \\ 0, & x \in B_\lambda \\ \in [0,x], & x \in C_\lambda \end{cases} \right\} \]

Since the densities are continuous, \( A_\lambda, B_\lambda \) are unions of open intervals (relative to [0, 1]) and \( C_\lambda \) is a closed set.

25
The corresponding Lagrangian to (3) is

\[ L(w(\cdot), \lambda) = E[x - w(x)|r] + \lambda(E_A[w(x)|r] - u) \]

In particular, the Lagrange multiplier method gives that if \( w \) solves the problem, there exists \( \lambda \in [m, M] \) such that

\[ \tilde{w} \in W_\lambda, \quad E_A[\tilde{w}(x)|r] = u. \]  

and that \( w = \tilde{w} \) almost everywhere. Conversely, any \( w \) satisfying (4) gives the principal the same payoff \( \Pi_r(u) \). Consequently, we will focus the characterization of optimal contracts within the sets \( W_\lambda \).

Define correspondence \( \Phi : [m, M] \to 2^{[0, E_A[x|r]]} \) by

\[ \Phi(\lambda) = \{ u : \exists w \in W_\lambda \text{ s.t. } E_A[w(x)|r] = u \} \]

For each \( \lambda \in [m, M] \), since \( W_\lambda \) is non-empty, \( \Phi(\lambda) \) is non-empty valued.

**Lemma A.1.** 1. Let \( m \leq \lambda < \lambda' \leq M \) and \( w \in W_\lambda, w' \in W_{\lambda'} \), then \( E_A[w(x)|r] < E_A[w'(x)|r] \).

2. For each \( \lambda \in [m, M] \), \( \Phi(\lambda) \) is a closed interval.

3. \( 0 \in \Phi(m), E_A[x|r] \in \Phi(M) \).

4. \( \Phi \) is upper-hemicontinuous.

**Proof of Lemma A.1.** 1. Let

\[ \bar{w}(x) = \begin{cases} x, & x \in A_{\lambda'} \\ 0, & x \notin A_{\lambda'} \end{cases} \]

and

\[ \underline{w}(x) = \begin{cases} x, & x \notin B_{\lambda} \\ 0, & x \in B_{\lambda} \end{cases} \]

Then \( \bar{w} \in W_{\lambda'}, \underline{w} \in W_\lambda, w'(x) \geq \bar{w}(x) \geq \underline{w}(x) \geq w(x) \) for all \( x \in [0, 1] \), hence \( E_A[w'(x)|r] \geq E_A[w(x)|r] \).
Let \( \epsilon \) be such that \( \lambda + \epsilon < \lambda' - \epsilon \). Then the set

\[
S = \{ x : (\lambda' - \epsilon) f_A(x|r) > f(x|r) \} \cap \{ x : (\lambda + \epsilon) f_A(x|r) < f(x|r) \}
\]

is open and \( \overline{w}(x) - \underline{w}(x) = x \) on \( S \). Since \( f_A(x|r) \) is continuous and everywhere positive,

\[
E_A[w'(x) - w(x)|r] \geq \int_S x f_A(x|r) dx > 0.
\]

2. Fix \( \lambda \). For each \( q \in [0, 1] \), define

\[
w_q(x) = \begin{cases} 
  x, & x \in A_\lambda \cup (C_\lambda \cap [0, q]) \\
  0, & \text{else}
\end{cases}
\]

Then \( w_q(x) \in W_\lambda \) for any \( q \in [0, 1] \). Furthermore, for any \( u \in \Phi(\lambda) \), \( E_A[w_0|r] \leq u \leq E_A[w_1|r] \). We now claim that the mapping \( q \mapsto E_A[w_q(x)|r] \) is continuous. To see this, note that for any \( q_2 > q_1 \),

\[
E_A[w_{q_2}(x) - w_{q_1}(x)|r] = E_A[x_1_{(q_1, q_2]} \cap C_\lambda(x)|r] \leq N(q_2 - q_1)
\]

where \( N = \max_{x \in [0, 1]} f_A(x|r) \). This shows that \( \Phi(\lambda) \) is the image of the interval \([0, 1]\) under a continuous mapping, hence is itself an interval.

Furthermore, let \( u \) be a limit point of the interval \( \Phi(\lambda) \). Let \( \{u_n\} \) be a sequence in \( \Phi(\lambda) \) that monotonically converges to \( u \). Choose \( q_n \in [0, 1] \) such that \( E_A[w_{q_n}|r] = u_n \). Then \( q_n \) is monotonic, with limit \( q \), and \( w_{q_n} = x_1_{A_\lambda \cup (C_\lambda \cap [0, q_n])} \) is a monotone sequence of bounded measurable functions. In particular, \( w = \lim_{n \to \infty} w_{q_n} = x_1_{A_\lambda \cup (C_\lambda \cap [0, q_n])}(x) \in W_\lambda \) and the Monotone Convergence Theorem implies \( u = \lim_{n \to \infty} E_A[w_{q_n}|r] = E_A[w|r] \in \Phi(\lambda) \). Hence \( \Phi(\lambda) \) is closed.

3. Let \( w \) be s.t. \( w(x) = 0 \) for all \( x \) and \( w' \) be s.t. \( w'(x) = x \) for all \( x \). The claim follows from \( w \in W_m \) and \( w' \in W_M \).

4. Let \( \lambda_n \to \lambda \) and \( u_n \in \Phi(\lambda_n) \) such that \( u_n \to u \). Without loss of generality (passing to subsequence if necessary), assume \( \lambda_n \) monotonically converges to \( \lambda \). For each \( n \), let
$w_n \in W_{\lambda_n}$ such that $E_A[w_n|r] = u_n$. In particular, $w_n(x)$ is a monotonic sequence because the set on which $w_n(x) = x$ gets either strictly larger or strictly smaller as $\lambda_n$ is increasing or decreasing. Hence the sequence has a measurable point-wise limit $w$. Since the Monotone Convergence Theorem gives $u = \lim_{n \to \infty} u_n = E_A[w|r]$, if $w \in W_\lambda$ then it is immediate that $u \in \Phi(\lambda)$.

To see that $w \in W_\lambda$, first observe that $w$ is measurable and $0 \leq w(x) \leq x$ because each $w_n$ is. It therefore suffices to show that $w(x) = x$ for $x \in A_\lambda$ and $w(x) = 0$ for $x \in B_\lambda$. Suppose $\lambda_n$ increases to $\lambda$. Then $w_n$ is an increasing sequence. If $x \in A_\lambda$ then there exists $n$ such that $x \in A_{\lambda_n}$. So $w(x) = \lim_{n \to \infty} w_n(x) = x$ because $\{w_n\}$ is an increasing sequence. If $x \in B_\lambda$, then $x \in B_{\lambda_n}$ for all $n$, and therefore $w(x) = \lim_{n \to \infty} w_n(x) = 0$.

Suppose $\lambda_n$ decreases to $\lambda$. Then $w_n$ is a decreasing sequence. If $x \in A_\lambda$, then $x \in A_{\lambda_n}$ for all $n$ and therefore $w(x) = \lim_{n \to } w_n(x) = x$. If $x \in B_\lambda$, then $x \in B_{\lambda_n}$ for some $n$.

Hence $w(x) = \lim_{n \to \infty} w_{\lambda_n}(x) = 0$ because $\{w_n\}$ is a decreasing sequence.

\[\square\]

**Lemma A.2.** The function $\lambda : [0, E_A[x|r]] \to [m, M]$ defined by $\lambda(u) = \Phi^{-1}(u)$ is a non-decreasing continuous function. Furthermore, for every $u \in [0, E_A[x|r]]$, any solution $(w, \lambda)$ to (4) satisfies $\lambda = \lambda(u)$ and $w \in W_{\lambda(u)}$.

**Proof.** Item 1 of **Lemma A.1** implies that $\Phi^{-1}(u)$ is at most a singleton and Item 3 implies that $\Phi^{-1}(0), \Phi^{-1}(E_A[x|r])$ are non-empty. To see that $\Phi^{-1}(u)$ is non-empty for $0 < u < E_A[x|r]$, let

\[L_u = \{\lambda : \max \Phi(\lambda) < u\}\]
\[R_u = \{\lambda : \min \Phi(\lambda) > u\}\]

**Lemma A.1** implies that $L_u, R_u$ and a common limit point $\lambda^*$ partition $[m, M]$ into left and right intervals. In particular, let $\lambda_n$ be an increasing sequence in $L_u$ converging to $\lambda^*$ and $\bar{\lambda}_n$ be a decreasing sequence in $R_u$ converging to $\lambda^*$. Define $u_n = \max \Phi(\lambda_n)$ and $\bar{u}_n = \min \Phi(\bar{\lambda}_n)$, which exist because $\Phi$ is closed. Moreover, \{u_n\}, \{\bar{u}_n\} are monotonic hence $u_n$ converges to a
limit $\bar{u}$, $\hat{u}_n$ converges to a limit $\bar{u}$, with $\bar{u} \leq u \leq \bar{u}$ and $\bar{u}, \bar{u} \in \Phi(\lambda^*)$ by upper-hemicontinuity. Since $\Phi(\lambda^*)$ is an interval, $u \in \Phi(\lambda^*)$. Therefore $\Phi^{-1}(u) = \lambda^*$.

That $\lambda(\cdot)$ is non-decreasing and continuous follow from Lemma A.1. If $(w, \lambda)$ solves (4), then by definition $u \in \Phi(\lambda)$ hence $\Phi^{-1}(u) = \lambda = \lambda(u)$. 

\[ \square \]

**Lemma A.3.** $\lambda(u)$ is directionally differentiable.

**Proof.** Let 

$$\phi(x) = \frac{f(x|r)}{f_A(x|r)},$$

which is well-defined and continuously differentiable for $x \in [0,1]$.

We show that $\lambda'(u+)$ exists. The argument for $\lambda'(u-)$ is identical.

We will use the following fact: For any $a$, there exists $\varepsilon$ such that for any $a < b < a + \varepsilon$, 

$$\{x : a < \phi(x) < b\}$$

is a disjoint union of $N < \infty$ open intervals $\{I_i\}$, where $N$ depends only on $\varepsilon$ (take some $I_i$ to be the empty set if necessary), and $\phi$ is strictly monotonic on each $I_i$.

It follows that for any $a < b < a + \varepsilon$, 

$$\{x : \phi(x) = b\}$$

is a set of finitely many points (equal to the number of the intervals $I_i$), hence it has zero Lebesgue measure.

Given $u \in \Phi(\lambda(u))$. Assume $u$ is in the interior of $\Phi(\lambda(u))$, then $\lambda'(u) = 0$. Assume $u$ is a limit point of $\Phi(\lambda)$ and that $u' \notin \Phi(\lambda)$ for every $u' > u$. Then $u = E_A[w|r]$ where $w$ is such that $w(x) = x$ if $x \notin B_{\lambda(u)}$, 0 otherwise. Since $\lambda(u)$ is continuous, there exists $\delta$ such that if $u < u' < u + \delta$, then $\lambda(u') - \lambda(u) < \varepsilon$, and $\lambda(u') > \lambda(u)$. For any such $u'$, $u' = E_A[w'|r]$ where $w'(x) = x$ whenever $x \in A_{\lambda(u')}$ and $w(x) = 0$ otherwise (since $\{x : \phi(x) = \lambda(u')\}$ is of measure zero). Therefore,

$$u' - u = E_A[w' - w|r] = \int_{\{x : \lambda(u) < \phi(x) < \lambda(u')\}} xf_A(x|r)dx$$

Let $\{x : \lambda(u) < \phi(x) < \lambda(u')\} = \bigcup_{i=1}^n(a_i, b_i)$, where the intervals are ordered: $a_1 \leq b_1 \leq a_2 \leq ... \leq b_n$. The mean value theorem for integrals gives $x_i(u') \in (a_i, b_i)$ such that

$$u' - u = \sum_{i=1}^N (b_i - a_i)x_i(u')f_A(x_i(u')|r)$$

29
For each \((a_i, b_i)\) such that \(b_i - a_i > 0\), the mean value theorem gives some \(y_i \in (\lambda(u), \lambda(u'))\) such that

\[ b_i - a_i = |(\phi^{-1})'(y_i)| (\lambda(u') - \lambda(u)) \]

Let

\[
m(u') = \min_{i: I_i \neq \emptyset} \inf_{k_i \in (a_i, b_i)} |(\phi^{-1})'(k_i)|
\]

\[
M(u') = \max_{i: I_i \neq \emptyset} \sup_{k_i \in (a_i, b_i)} |(\phi^{-1})'(k_i)|
\]

Let

\[
K(u') = \sum_{i: I_i \neq \emptyset} x_i(u') f_A(x_i(u)) r
\]

Then we have

\[
m(u') K(u') (\lambda(u') - \lambda(u)) \leq u' - u \leq M(u') (K(u') (\lambda(u') - \lambda(u))
\]

Therefore

\[
\frac{1}{M(u') K(u')} \leq \frac{\lambda(u') - \lambda(u)}{u' - u} \leq \frac{1}{m(u') K(u')}
\]

Note that \(m(u')\) is bounded away from zero and that \(\sup_{u' \in (u, u + \delta)} K(u') < \infty\). Since \(\phi\) is continuously differentiable, \(\lim_{u' \to u} K(u') (M(u') - m(u')) = 0\), this implies \(\lambda(u)\) has a right derivative.

The next lemma is a version of the Implicit Function Theorem that we use to establish the Envelope Theorem.

**Lemma A.4.** Suppose \(F(\beta, u) = 0\) defines an implicit function \(u(\cdot)\) at a neighborhood of \(\beta\) that is strictly increasing and continuous. Suppose that either

1. \(F(\beta, u)\) is differentiable in \(\beta\), \(\partial F/\partial \beta\) is continuous in \(u\), and \(F(\beta, u)\) is directionally differentiable in \(u\), or

2. \(F(\beta, u)\) is directionally differentiable in \(\beta\), differentiable in \(u\) and \(\partial F/\partial u\) is continuous in \(\beta\).
Then $u(\cdot)$ is directionally differentiable at $\beta$.

**Proof.** Let $\Delta \beta > 0$. Note that

$$F(\beta + \Delta \beta, u(\beta + \Delta \beta)) - F(\beta, u(\beta)) = 0$$

Since $u(\cdot)$ is strictly increasing, we can write

$$\frac{F(\beta + \Delta \beta, u(\beta + \Delta \beta)) - F(\beta + \Delta \beta, u(\beta))}{\Delta \beta} + \frac{F(\beta, u(\beta + \Delta \beta)) - F(\beta, u(\beta))}{u(\beta + \Delta \beta) - u(\beta)} \frac{u(\beta + \Delta \beta) - u(\beta)}{\Delta \beta} = 0$$

(5)

$$\frac{F(\beta + \Delta \beta, u(\beta + \Delta \beta)) - F(\beta + \Delta \beta, u(\beta))}{u(\beta + \Delta \beta) - u(\beta)} + \frac{F(\beta, u(\beta + \Delta \beta)) - F(\beta, u(\beta))}{\Delta \beta} = 0$$

(6)

In Case 1, since $\partial F/\partial \beta$ is continuous in $u$, letting $\Delta \beta \to 0$ in (5) we can obtain

$$\frac{\partial F(\beta, u(\beta))}{\partial \beta} + \frac{\partial F(\beta, u(\beta)+)}{\partial u} u'(\beta+) = 0$$

Hence $u(\cdot)$ is right-differentiable. Performing the same argument with $\Delta \beta < 0$ shows that $u(\cdot)$ is left-differentiable.

In Case 2, since $\partial F/\partial u$ is continuous in $\beta$, letting $\Delta \beta \to 0$ in (6) we can obtain

$$\frac{\partial F(\beta, u(\beta))}{\partial u} u'(\beta+) + \frac{\partial F(\beta, u(\beta))}{\partial \beta} = 0$$

Hence $u(\cdot)$ is right-differentiable. Performing the same argument with $\Delta \beta < 0$ shows $u(\cdot)$ is left-differentiable.

**Lemma A.5.** Consider a two player normal form game such that player $i$’s payoff is

$$p(u_i, u_{-i}; \beta)\Pi(u_i),$$

where player $i$ chooses action $u_i \in [0, \bar{u}]$. Suppose $\Pi$ is strictly decreasing, continuously differentiable, weakly concave, $\Pi(0) > 0$ and $\Pi(\bar{u}) = 0$, $\Pi''(u+), \Pi''(u-)$ exist, and $p$ satisfies Assumption 1. Then

1. For each $\beta$ there exists a unique symmetric Nash equilibrium $u(\beta)$ such that $u(\cdot)$ is continuously increasing, $u(0) = 0$ and $\lim_{\beta \to \infty} u(\beta) = \bar{u}$. 

31
2. $u(\cdot)$ is directionally differentiable. When $\Pi(u)$ is linear, $u(\cdot)$ is differentiable at any $\beta$ such that $u(\beta) > 0$.

3. Let

$$E\Pi(\beta) = \max_{u_i \in [0,\overline{u}]} p(u_i, \bar{u}(\beta), \beta)\Pi(u_i)$$

where $\bar{u}(\cdot)$ is any continuously increasing and directionally differentiable function. Let $u^*(\beta)$ be the optimal $u_i$ for each $\beta$. If either

a. $\bar{u}$ is directionally differentiable and $\Pi$ is linear, or

b. $\bar{u}$ is differentiable, or

c. $u^*(\cdot) = \bar{u}(\cdot)$ and they are directionally differentiable,

Then at any $\beta$ s.t. $u^*(\beta) > 0$,

$$E\Pi'(\beta+) = p_2(u^*(\beta), \bar{u}(\beta), \beta)\bar{u}'(\beta+) + p_3(u^*(\beta), \bar{u}(\beta), \beta)\Pi(u^*(\beta))$$
$$E\Pi'(\beta-) = p_2(u^*(\beta), \bar{u}(\beta), \beta)\bar{u}'(\beta-) + p_3(u^*(\beta), \bar{u}(\beta), \beta)\Pi(u^*(\beta))$$  \hspace{1cm} (7)

**Proof of Lemma A.5.** The first order condition for symmetric pure strategy Nash equilibrium is

$$\frac{p_1(u, u; \beta)}{p(u, u; \beta)} \leq -\frac{\Pi'(u)}{\Pi(u)}, u \geq 0, \text{with complementary slackness}$$  \hspace{1cm} (8)

By Assumption 1.3, the left-hand side is non-increasing in $u$. Since $\Pi$ is weakly concave and strictly decreasing, the right-hand side is strictly increasing in $u$. Both sides are continuous and the right-hand side ranges from 0 to $\infty$ as $u$ increases from 0 to $\overline{u}$. Therefore, (8) has a unique solution, denoted by $u(\beta)$. By Assumption 1.3, $p$ is log-concave in $u_1$, so $u(\beta)$ is indeed a mutual best response.

That $u(\beta)$ increases from zero to $\overline{u}$ as $\beta \to \infty$ is a consequence of Assumption 1.3, that the left-hand side of (8) increases from zero to infinity as $\beta \to \infty$ for each $u$, and that the right-hand side increases to infinity as $u \to \overline{u}$. The continuity of $u(\cdot)$ follows from the continuity of $p_1/p$ in $\beta$. This proves item 1.
Next, we show item 2. Let

\[ F(\beta, u) = \frac{p_1(u, u; \beta)}{p(u, u; \beta)} + \frac{\Pi'(u)}{\Pi(u)}, \]

which is directionally differentiable in \( u \) and differentiable in \( \beta \). Furthermore, \( \partial F / \partial \beta \) is continuous in \( u \) since \( p \) is \( C^2 \).

Suppose that \( u(\beta) > 0 \), then (8) holds with equality and \( u(\cdot) \) is strictly increasing at \( \beta \). Applying Case 1 of Lemma A.4 shows that \( u'(\beta^+), u'(\beta^-) \) exist.

Suppose that \( u(\beta) = 0 \) and that \( u(\beta + \epsilon) = 0 \) for some \( \epsilon > 0 \). Then there is a neighborhood of \( \beta \) such that \( u(\cdot) = 0 \). So \( u(\beta) \) is differentiable. Suppose that \( u(\beta) = 0 \) and that \( u(\beta + \epsilon) > 0 \) for all \( \epsilon > 0 \). The continuity of \( F \) shows that (8) holds with equality. The same argument in Lemma A.4 shows that \( u \) is right-differentiable at \( \beta \). Since \( u(\beta') = 0 \) for all \( \beta' < \beta \), \( u \) is left-differentiable at \( \beta \) as well. When \( \Pi \) is linear, the FOC that \( u(\beta) > 0 \) satisfies is a \( C^1 \) function, hence the standard Implicit Function Theorem implies that \( u(\cdot) \) is differentiable at \( \beta \).

Finally we show item 3, which is a version of the Envelope Theorem when the maximizer is directionally differentiable.

Since the objective function is differentiable, \( u^*(\beta) \) satisfies the FOC

\[ \frac{p_1(u, \tilde{u}(\beta); \beta)}{p(u, \tilde{u}(\beta); \beta)} \leq \frac{\Pi'(u)}{\Pi(u)}, \quad u \geq 0, \quad \text{with complementary slackness} \]

Suppose \( u^*(\beta) > 0 \). Then the FOC is satisfied at a neighborhood of \( \beta \). In Case a and b an application of Lemma A.4 shows that \( u^* \) is directionally differentiable.

Take \( \epsilon > 0 \). We have

\[ E\Pi(\beta + \epsilon) - E\Pi(\beta) = p(u(\epsilon), \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u(\epsilon)) - p(u, \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u(\epsilon)) + p(u, \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u) - p(u(\epsilon), \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u(\epsilon)) + p(u(\epsilon), \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u(\epsilon)) - p(u(\epsilon), \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u) + p(u(\epsilon), \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u) - p(u, \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u) + p(u, \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u) - p(u, \tilde{u}(\beta + \epsilon); \beta + \epsilon)\Pi(u) \]
A similar argument to item 1 using log-concavity of $p$ shows that $u(\epsilon) > u > 0$. Hence

\[
E\Pi(\beta + \epsilon) - E\Pi(\beta) = \frac{p(u(\epsilon), \tilde{u}(\beta + \epsilon); \beta + \epsilon) - p(u, \tilde{u}(\beta + \epsilon); \beta + \epsilon)}{u(\epsilon) - u} \frac{u(\epsilon) - u}{\epsilon} \Pi(u(\epsilon))
\]

\[
+ p(u, \tilde{u}(\beta + \epsilon); \beta + \epsilon) \left( \frac{\Pi(u(\epsilon)) - \Pi(u) u(\epsilon) - u}{u(\epsilon) - u} \frac{u(\epsilon) - u}{\epsilon} \right) \Pi(u)
\]

\[
+ \left( \frac{p(u, \tilde{u}(\beta + \epsilon); \beta + \epsilon) - p(u, \tilde{u}(\beta); \beta + \epsilon)}{\tilde{u}(\beta + \epsilon) - \tilde{u}(\beta)} \frac{\tilde{u}(\beta + \epsilon) - \tilde{u}(\beta)}{\epsilon} \right) \Pi(u)
\]

Taking $\epsilon \to 0$ and noting that $p$ is $C^2$, we get

\[
E\Pi'(\beta+) = p_1\Pi'(u^*(\beta))(u^*)'(\beta+) + p_2\Pi'(u^*(\beta))\tilde{u}'(\beta+) + p_3\Pi(u^*(\beta))
\]

\[
= p_2\Pi'(u^*(\beta))(\tilde{u}')(\beta+) + p_3\Pi(u^*(\beta))
\]

since the first two terms combine to zero by the FOC $u^*(\beta) > 0$ must satisfy. Taking $\epsilon < 0$ and performing the same argument takes care of $E\Pi'(\beta-)$. 

\[\square\]

**Corollary A.1.** Let $u_s(\cdot), u_r(\cdot)$ be the implicit functions derived in Lemma A.5 when $\Pi_s, \Pi_r$ are in place of $\Pi$, respectively. Then $u_s(\cdot), u_r(\cdot)$ are directionally differentiable.

**Proof.** Since $\Pi_s(u)$ is linear, the case for $u_s(\cdot)$ follows straightforwardly from item 3 of Lemma A.5.

Applying Corollary 5 of Milgrom & Segal (2002) to (3), we get $\Pi'(u) = -\lambda(u)$. It then follows from Lemma A.3 that $\Pi'(u+), \Pi'(u-)$ exist. Hence Lemma A.5 is again applicable. 

\[\square\]

In particular, the above corollary allows us to apply Lemma A.5 to (2).

**Lemma A.6.** For all $0 \leq u < E[x|s],

\[
\frac{1}{\Pi_s(u)} > \frac{-\Pi'(u)}{\Pi_r(u)}.
\]

In addition, $u_s(\beta) < u_r(\beta)$ for all $\beta$ such that $u_s(\beta) > 0$. 

34
Proof. The first-order conditions \( u_s, u_r \) must satisfy are

\[
\frac{p_1(u, u; \beta)}{p(u, u; \beta)} \leq \frac{1}{\Pi_s(u)}, \ u \geq 0, \ \text{with complementary slackness}
\]

\[
\frac{p_1(u, u; \beta)}{p(u, u; \beta)} \leq -\frac{\Pi'_r(u)}{\Pi_r(u)}, \ u \geq 0, \ \text{with complementary slackness}
\]

Hence (9) implies \( u_s(\beta) < u_r(\beta) \) whenever \( u_s(\beta) > 0 \).

To see (9), note that \( \Pi_s(u) = E[x|s] - u, \Pi_r(u) = E[x - w(x)|r] \) where \( w(x) \in \mathcal{W}_{\lambda(u)} \), and that \( -\Pi'_r(u) = \lambda(u) \). Furthermore, the proof of Lemma A.1 Item 2 implies that we can choose \( w(x) \) to be of the form \( w(x) = x1_{A_{\lambda(u)} \cup (C_{\lambda(u)} \cap [0, q])}(x) \) for some \( q \in [0, 1] \). The set \( S = A_{\lambda(u)} \cup (C_{\lambda(u)} \cap [0, q]) \) on which \( w(x) = x \) is a measurable subset of \( \{ x : \lambda(u) \geq f(x)r/f_A(x)r \} \). Let \( S^c \) denote its complement. We also write \( E[w(x)|r] = \int_S xf(x)r dx := E[x|r, S] \). \( E_A[x|r, S^c], E[x|r, S^c] \) are similarly defined. In particular, \( E_A[x|r, S^c] = E_A[x|r] - u \) because the IR constraint is tight.

Suppose to the contrary,

\[
\frac{1}{\Pi_s(u)} \leq -\frac{\Pi'_r(u)}{\Pi_r(u)}
\]

Then

\[
E[x - w(x)|r] \leq \lambda(u)(E[x|s] - u)
\]

By the definition of \( S \) and \( \lambda(u) \),

\[
xf_A(x|r)(E[x|r] - E[x|r, S]) \leq xf(x)r(E[x|s] - u), \ \forall x \in S^c.
\]

Integrate the above equation with respect to \( x \) over \( S^c \) to obtain

\[
E_A[x|r, S^c](E[x|r] - E[x|r, S]) \leq E[x|r, S^c](E[x|s] - u)
\]

Substitute in \( E_A[x|r, S^c] = E_A[x|r] - u, E[x|r, S^c] = E[x|r] - E[x|r, S] \) and carry out a cancellation to get

\[
E_A[x|r] \leq E[x|s],
\]

a contradiction to our assumption of expected value dominance. \( \square \)

The following lemma proves some properties of the optimal deviations \( u_s^e, u_r^e \).
Lemma A.7.  
1. $u^*_r(\cdot)$ is continuous. $u^*_r(0) = 0$.

2. $u^*_r(\cdot)$ is continuous and increasing. $\lim_{\beta \to \infty} u^*_r(\beta) = E[x|s]$. In particular, there exists $\beta$ s.t. $u^*_r(\beta) > u^*$.

Proof. The first order condition that $u = u^*_r(\beta)$ must satisfy is

$$\frac{p_1(u, u_r(\beta); \beta)}{p(u, u_r(\beta); \beta)} \leq -\frac{\Pi'_r(u)}{\Pi_r(u)}, u \geq 0,$$

which defines a continuous function because both sides are continuous. When $\beta = 0$, Assumption 1.3 says that the left-hand side is zero. This forces $u^*_r(0) = 0$.

To prove the remaining properties of $u^*_r$, first we claim that for all $\beta$,

$$u^*_r(\beta) \leq u_r(\beta).$$

To see this, simply note that $u^*_r(\beta)$ and $u_r(\beta)$ satisfy the FOCs

$$\frac{p_1(u, u_r(\beta); \beta)}{p(u, u_r(\beta); \beta)} \leq -\frac{\Pi'_r(u)}{\Pi_r(u)}, u \geq 0,$$

with complementary slackness.

By Lemma A.6, $1/\Pi_s(u) > -\Pi'_r(u)/\Pi_r(u)$ for all $u > 0$. Since $p$ is log-concave in $u_1$, it implies $u^*_r(\beta) \leq u_r(\beta)$ for all $\beta \geq 0$.

To see that $u^*_r(\cdot)$ is increasing, note that $p_1(u_1, u_2; \beta)/p(u_1, u_2; \beta)$ is weakly increasing in $\beta$ for all $u_1 \leq u_2$ and weakly increasing in $u_2$. Let $\beta > \beta'$. Suppose to the contrary that $u^*_r(\beta) < u^*_r(\beta')$. Since $u_r(\beta) \geq u_r(\beta')$ and $u^*_r(\beta) \leq u_r(\beta)$, we have

$$\frac{p_1(u^*_r(\beta), u_r(\beta); \beta)}{p(u^*_r(\beta), u_r(\beta); \beta)} \geq \frac{p_1(u^*_r(\beta), u_r(\beta'); \beta')}{p(u^*_r(\beta'), u_r(\beta'); \beta')} > \frac{p_1(u^*_r(\beta'), u_r(\beta); \beta)}{p(u^*_r(\beta'), u_r(\beta'); \beta')},$$

where in the last inequality we used that $p$ is log-concave in $u_1$. However, this violates the FOCs $u^*_r(\beta)$ and $u^*_r(\beta')$ must satisfy because $1/\Pi_s(u^*_r(\beta)) < 1/\Pi_s(u^*_r(\beta'))$.

To see that $\lim_{\beta \to \infty} u^*_r(\beta) = E[x|s]$, note that log-concavity of $p$ with respect to $u_1$ implies that

$$\frac{p_1(u^*_r(\beta), u_r(\beta); \beta)}{p(u^*_r(\beta), u_r(\beta); \beta)} \geq \frac{p_1(u_r(\beta), u_r(\beta); \beta)}{p(u_r(\beta), u_r(\beta); \beta)}$$

and the latter tends to infinity as $\beta \to \infty$. The FOC then implies $u^*_r(\beta) \to \infty$. □
Proof of Lemma 3.2. By Proposition 3.1, $E\Pi_s(\beta) > E\Pi_s^r(\beta)$ whenever $u_s^*(\beta) < u^*$ and $E\Pi_s(\beta) < E\Pi_s^r(\beta)$ whenever $u_s(\beta) > u^*$. (The principal can simply replace a risky contract with a safe contract and keep the promised utility to the agent fixed, and vice versa.) By Lemma A.5, there is a $\beta$ such that $u_s(\beta) > u^*$, and by Lemma A.7 there is a $\beta$ such that that $u_s^*(\beta) < u^*$. Consequently, $E\Pi_s$ and $E\Pi_s^r$ cross at least once within $(0, \infty)$.

Let $\beta_s$ be a point such that $E\Pi_s(\beta_s) = E\Pi_s^r(\beta_s)$. Then $u_s(\beta_s) \leq u^* \leq u_s^*(\beta_s)$. Moreover, since $\Pi_s^r(u^*) > \Pi_s^r(u^*) = -1$, at least one of the inequalities is strict. (Else one of $u_s, u_s^*$ will violate its FOC)

By Lemma A.5,

$$E\Pi_s^r(\beta_s+) = p_2(u_s(\beta_s), u_s(\beta_s); \beta_s)u_s^r(\beta_s+)) + p_3(u_s(\beta_s), u_s(\beta_s); \beta_s))\Pi_s(u_s(\beta_s)) \leq 0$$

$$(E\Pi_s^r)'(\beta_s+) = p_2(u_s^r(\beta_s), u_s(\beta_s); \beta_s)u_s^r(\beta_s+)) + p_3(u_s^r(\beta_s), u_s(\beta_s); \beta_s))\Pi_r(u_s^r(\beta_s)) \leq 0$$

Since $E\Pi_s(\beta_s) = p(u_s, u_s; \beta_s)\Pi_s(u_s) = p(u_s^r, u_s; \beta_s)\Pi_r(u_s^r) = E\Pi_s^r(\beta_s)$, after substitution and cancellation we obtain

$$E\Pi_s^r(\beta_s+) < (E\Pi_s^r)'(\beta_s+)$$

$$\iff \frac{p_2(u_s(\beta_s), u_s(\beta_s); \beta_s)}{p(u_s, u_s; \beta_s)} u_s^r(\beta_s+) + \frac{p_3(u_s(\beta_s), u_s(\beta_s); \beta_s)}{p(u_s, u_s; \beta_s)} < \frac{p_2(u_s^r(\beta_s), u_s^r(\beta_s); \beta_s)}{p(u_s^r, u_s^r; \beta_s)} u_s^r(\beta_s+) + \frac{p_3(u_s^r(\beta_s), u_s^r(\beta_s); \beta_s)}{p(u_s^r, u_s^r; \beta_s)}$$

which follows from the assumptions about $p$. An identical argument shows that this inequality also holds for left-handed derivatives. Therefore, whenever $E\Pi_s$ and $E\Pi_s^r$ cross, the former has a more negative slope in both the left and the right directions. Thus $E\Pi_s^r$ crosses $E\Pi_r$ exactly once, from below.

Now we prove the result for $E\Pi_r(\beta)$. By Proposition 3.1, $E\Pi_r(\beta) < E\Pi_r^s(\beta)$ whenever $u_r(\beta) < u^*$, and $E\Pi_r(\beta) > E\Pi_r^s(\beta)$ whenever $u_r^s(\beta) > u^*$. By Lemma A.5, there is a $\beta$ such that $u_r(\beta) < u^*$ and by Lemma A.7 there is a $\beta$ such that $u_r^s(\beta) > u^*$. Therefore they cross at least once.

Choose any $\beta_r$, such that

$$E\Pi_r(\beta_r) = E\Pi_r^s(\beta_r).$$

---

19If $u_s(\beta_s) > u^*$, then it has to be $E\Pi_s^r(\beta_s) > E\Pi_s(\beta_s)$ by Proposition 3.1. Similarly, if $u_s^*(\beta_s) < u^*$, then it has to be $E\Pi_s^r(\beta_s) < E\Pi_s(\beta_s)$. 

37
Then by previous discussions \(0 < \beta_s < \infty\). Moreover, \(u^*_r(\beta_r) \leq u^* \leq u_r(\beta_r)\). A similar argument as before shows that \(u^*_r(\beta_r) < u_r(\beta_r)\). By Lemma A.5,

\[
E\Pi_r'(\beta_r+) = (p_2(u_r(\beta_r), u_r(\beta_r); \beta_r)u_r'(\beta_r+) + p_3(u_r(\beta_r), u_r(\beta_r); \beta_r))\Pi_r(u_r(\beta_r)) \leq 0
\]

\[
(\Pi^s_r)'(\beta_r+) = (p_2(u^*_s(\beta_r), u_r(\beta_r); \beta_r)u_r'(\beta_r+) + p_3(u^*_s(\beta_r), u_r(\beta_r); \beta_r))\Pi_s(u^*_r(\beta_r)) \leq 0
\]

Since \(E\Pi_r(\beta_r) = p(u_r, u_r; \beta_r)\Pi_r(u_r) = p(u^*_r, u_r; \beta_r)\Pi_s(u^*_r) = E\Pi^s_r(\beta_r)\), after substitution and cancellation we obtain

\[
E\Pi_r'(\beta_r+) > (\Pi^s_r)'(\beta_r+)
\]

\[
\iff \frac{p_2(u_r(\beta_r), u_r(\beta_r); \beta_r)}{p(u_r(\beta_r), u_r(\beta_r); \beta_r)}u_r'(\beta_r+) + \frac{p_3(u_r(\beta_r), u_r(\beta_r); \beta_r)}{p(u_r(\beta_r), u_r(\beta_r); \beta_r)} > \frac{p_2(u^*_s(\beta_r), u_r(\beta_r); \beta_r)}{p(u^*_s(\beta_r), u_r(\beta_r); \beta_r)}u_r'(\beta_r+) + \frac{p_3(u^*_s(\beta_r), u_r(\beta_r); \beta_r)}{p(u^*_s(\beta_r), u_r(\beta_r); \beta_r)}
\]

which again follows from the assumptions about \(p\). An identical argument shows the inequality for the left-derivatives.

Consequently, whenever \(E\Pi^s_r\) and \(E\Pi_r\) crosses, the former has a more negative left and right derivative. So \(E\Pi^s_r\) crosses \(E\Pi_r\) exactly once, from above.

Finally, we prove that \(\beta_r < \beta_s\). To do this, we show that at \(\beta_s\), the principals strictly prefer not to deviate from the risky equilibrium. By the definition of \(\beta_s\), we have \(E\Pi_s(\beta_s) = E\Pi^s_s(\beta_s)\). Fix \(\beta\) at \(\beta_s\), and consider the optimal best response of the principal with respect to \(u_2\). Let \(\tilde{u}_a(u_2)\) and \(V_a(u_2), a \in \{s, r\}\), denote the solution and value function of

\[
\max_u p(u, u_2; \beta_s)\Pi_a(u).
\]

Note that for \(u_2 = u_s(\beta_s)\), we have \(\tilde{u}_s(u_2) = u_s(\beta_s), \tilde{u}_r(u_2) = u^*_r(\beta_s)\), and that

\[
V_s(u_2) = E\Pi_s(\beta_s) = E\Pi^s_s(\beta_s) = V_r(u_2).
\]

Note also that for \(u_2 = u_r(\beta_s)\), we have \(\tilde{u}_s(u_2) = u^*_r(\beta_s), \tilde{u}_r(u_2) = u_r(\beta_s)\), and that

\[
V_s(u_2) = E\Pi^s_r(\beta_s), V_r(u_2) = E\Pi_r(\beta_s).
\]

---

\(^{20}\)If \(u^*_r(\beta_r) > u^*\), then it has to be \(E\Pi^s_r(\beta_r) < E\Pi_r(\beta_r)\) by Proposition 3.1. Similarly, if \(u_r(\beta_r) < u^*\), then it has to be \(E\Pi^s_r(\beta_r) > E\Pi_r(\beta_r)\).
We aim to show that whenever \( V_s, V_r \) cross each other, \( V_s \) must cross \( V_r \) from above. This would imply that \( V_s \) stays below \( V_r \) at all \( u_2 \) right-ward to the crossing point. In particular, by Lemma A.6 \( u_r(\beta) > u_s(\beta) \), so we would have \( E\Pi_r(\beta) < E\Pi_s(\beta) \), and this implies that \( \beta_r < \beta_s \).

The standard envelope theorem argument implies that

\[
V_s'(u_2) = p_2(u_s(\beta_s), u_s(\beta_s); \beta_s)\Pi_s(u_s(\beta_s)) \leq 0
\]

\[
V_r'(u_2) = p_2(u_r^*(\beta_s), u_s(\beta_s); \beta_s)\Pi_r(u_r^*(\beta_s)) \leq 0
\]

When \( V_s(u_2) = V_r(u_2) \), we have

\[
p(u_s(\beta_s), u_s(\beta_s); \beta_s)\Pi_s(u_s(\beta_s)) = p(u_r^*(\beta_s), u_r(u_2); \beta_s)\Pi_r(u_r^*(\beta_s))
\]

Therefore, at the crossing point \( u_2 \), \( V_s'(u_2) < V_r'(u_2) \) if and only if

\[
\frac{p_2(\tilde{u}_s(u_2), u_2; \beta_s)}{p(\tilde{u}_s(u_2), u_2; \beta_s)} < \frac{p_2(\tilde{u}_r(u_2), u_2; \beta_s)}{p(\tilde{u}_r(u_2), u_2; \beta_s)}
\]

(10)

Since \( \tilde{u}_s, \tilde{u}_r \) satisfy the first-order conditions

\[
\frac{p_1(u, u_2; \beta_s)}{p(u, u_2; \beta_s)} = \frac{1}{\Pi_s(u)}
\]

\[
\frac{p_1(u, u_2; \beta_s)}{p(u, u_2; \beta_s)} = -\frac{\Pi_r'(u)}{\Pi_r(u)}
\]

when they are positive, Lemma A.6 implies that \( \tilde{u}_r(u_2) > \tilde{u}_s(u_2) \) whenever \( \tilde{u}_s(u_2) > 0 \). Now (10) follows from our assumption that \( p_2/p \) is increasing in \( u_1 \).

To prove Theorem 2, we establish the following lemma governing how fast \( u_s, u_r \) increase with respect to \( \alpha \).

**Lemma A.8.** Under Assumption 2, for all \((\beta, \alpha)\) such that \( u_s(\beta, \alpha) > 0, u_r(\beta, \alpha) > 0 \),

\[
\frac{\partial u_s(\beta, \alpha)}{\partial \alpha} \geq \frac{u_s(\beta, \alpha)}{\alpha} \tag{11}
\]

\[
\frac{\partial u_r(\beta, \alpha)}{\partial \alpha} \geq \frac{u_r(\beta, \alpha)}{\alpha} \tag{12}
\]

39
In particular, for every $\beta > 0$ there exists $\bar{\alpha} > 0$ such that

$$
u_s(\beta, \alpha) = \begin{cases} 0, & \alpha \leq \bar{\alpha} \\ (\alpha - \bar{\alpha})E[x|s] & \alpha > \bar{\alpha} \end{cases} \tag{13}$$

**Proof.** We first show (13), which then implies (11). To see this, note that the by assumption $p_1/p > 0$ when $\beta > 0$. If $u_s(\beta, \alpha) = 0$ for every $\alpha$, then as $\alpha$ becomes large we would have

$$\frac{p_1(0, 0; \beta)}{p(0, 0; \beta)} > \frac{1}{\alpha E[x|s]}$$

which violates the first-order condition. Let $\bar{\alpha}$ be the minimum $\alpha$ such that $u_s(\beta, \bar{\alpha}) = 0$. By Assumption 2, $p_1(u, u; \beta)/p(u, u; \beta) = \xi(\beta)$. For any $\alpha > \bar{\alpha}$, since $u_s$ is increasing in $\alpha$, $u_s(\beta, \alpha) > 0$. By the FOC we have

$$\frac{1}{\alpha E[x|s] - u_s(\beta, \alpha)} = \xi(\beta) = \frac{1}{\bar{\alpha} E[x|s]}$$

Therefore, for $\alpha > \bar{\alpha}$,

$$u_s(\beta, \alpha) = (\alpha - \bar{\alpha})E[x|s].$$

and (13) follows.

Now we prove (12). For notational simplicity, write $u_r(\beta, \alpha) = u_r(\alpha)$. Whenever $u_r(\alpha) > 0$, the FOC is satisfied:

$$\frac{\lambda(u_r(\alpha), \alpha)}{E[\alpha x|r] - \int_{S(u_r(\alpha), \alpha)} \alpha x f_A(x|r)dx} = \xi(\beta) \tag{14}$$

where

$$S(u_r(\alpha), \alpha) = A_{\lambda(u_r(\alpha), \alpha)} \cup (C_{\lambda(u_r(\alpha), \alpha)} \cap [0, q(u_r(\alpha), \alpha)])$$

for some appropriately chosen $q(u_r(\alpha), \alpha) \in [0, 1]$ and the sets $A_{\lambda}, C_{\lambda}, [0, q(u_r(\alpha), \alpha)]$ are as defined in the beginning of Appendix A.1 and Lemma A.1 (with the modification of scaling up $x$ to $\alpha x$).

Since $u_r(\alpha)$ is increasing in $\alpha$, (14) continues to hold for larger $\alpha$.

Assume to the contrary that at some $(\beta, \alpha)$ such that $u_r(\beta, \alpha) > 0$,

$$\frac{\partial u_r(\beta, \alpha)}{\partial \alpha} < \frac{u_r(\beta, \alpha)}{\alpha}. \tag{15}$$
Since
\[ \frac{u_r(\alpha)}{\alpha} = \int_{S(u_r(\alpha), \alpha)} x f_A(x|r) dx \] (16)
and
\[ \frac{\partial u_r(\alpha)}{\partial \alpha} < 0 \Leftrightarrow \frac{\partial u_r(\beta, \alpha)}{\partial \alpha} < \frac{u_r(\beta, \alpha)}{\alpha}, \]
(15) implies that there exists an \( \epsilon > 0 \) such that for all \( \alpha < \alpha' < \alpha + \epsilon \)
\[ \int_{S(u_r(\alpha'), \alpha')} x f_A(x|r) dx < \int_{S(u_r(\alpha), \alpha)} x f_A(x|r) dx \] (17)
Thus for all \( \alpha' \in (\alpha, \alpha + \epsilon) \), \( \lambda(u_r(\alpha'), \alpha') \leq \lambda(u_r(\alpha), \alpha) \), as otherwise there will be an open set contained in \( S(u_r(\alpha'), \alpha') \cap S(u_r(\alpha), \alpha)^c \) (by the continuity of \( f(x|r)/f_A(x|r) \)), leading to a contradiction of (17).
This means that the numerator of the LHS of (14) weakly decreases, while the denominator strictly increases, when \( \alpha \) increases to any \( \alpha' \in (\alpha, \alpha + \epsilon) \). This is a contradiction to (14) being satisfied for all \( \alpha' > \alpha \). Therefore (12) is true. \( \square \)

**Proof of Theorem 2.** The existence of cutoffs \( \beta_r(\alpha) < \beta_s(\alpha) \) follows from identical steps as in Theorem 1.

We first show that \( \beta_s(\cdot) \) is decreasing. It suffices to show that whenever \( E\Pi_s(\beta, \alpha) = E\Pi_r(\beta, \alpha) \),
\[ \frac{\partial E\Pi_s}{\partial \alpha} < \frac{\partial E\Pi_r}{\partial \alpha}, \] (18)
where \( E\Pi_s(\beta, \alpha) = p(u_s(\beta, \alpha), u_s(\beta, \alpha); \beta)\Pi_s(u_s(\beta, \alpha), \alpha) \) and
\[ E\Pi_r(\beta, \alpha) = \max_{u \in [0, 1] S(u, \alpha)} p(u, u_s(\beta, \alpha); \beta)\Pi_r(u, \alpha). \]
This single-crossing from below condition implies that as \( \alpha \) increases, deviations to risky contracts from a safe equilibrium become more profitable.

Given \( u \) and \( \alpha \), let \( w_q \) be the optimal wage scheme of the form in Lemma A.1 and \( S(u, \alpha) \) be the set on which \( w_q(x) = \alpha x \). Further, for a measurable set \( S \subset [0, 1] \), we let \( E\Pi_r, E\Pi_s \), which will be used later, are defined in the same fashion.
\[ E[x|S] = \int_S xf(x)dx. \] Then using this notation

\[
\frac{\partial \Pi_s(u, \alpha)}{\partial \alpha} = E[x|s] \\
\frac{\partial \Pi_r(u, \alpha)}{\partial \alpha} = E[x|r] - \int_{S(u, \alpha)} xf(x|r)dx + \lambda(u, \alpha) \int_{S(u, \alpha)} xf_A(x|r)dx \\
= E[x|r, S^c(u, \alpha)] + \lambda(u, \alpha) E_A[x|r, S(u, \alpha)]
\]

where the second equality follows from applying the Envelope Theorem to the corresponding Lagrangian.

Suppose \( E\Pi_s(\beta, \alpha) = E\Pi_r^s(\beta, \alpha) \). That is,

\[
p(u_s(\beta, \alpha), u_s^r(\beta, \alpha); \beta)\Pi_s(u_s(\beta, \alpha), \alpha) = p(u_s^r(\beta, \alpha), u_s(\beta, \alpha); \beta)\Pi_r(u_s^r(\beta, \alpha), \alpha). \tag{19}
\]

Suppose that \( \alpha \neq \bar{\alpha} \) at the \((\beta, \alpha)\) such that (19) holds. For notational convenience, shorten \( u_s(\alpha) = u_s(\beta, \alpha), u_s^r(\alpha) = u_s^r(\beta, \alpha) \). We can prove using identical steps as Lemma A.5 that the right derivatives of the two sides of (19) with respect to \( \alpha \) are given by:

\[
E\Pi'_s(\alpha+) = p_2(u_s, u_s^r; \beta)u_s'(\alpha+)\Pi_s(u_s, \alpha) + p(u_s, u_s; \beta)E[x|s], \tag{20}
\]

\[
(E\Pi'_r)^{\prime}(\alpha+) = p_2(u_s^r, u_s^r; \beta)u_s'(\alpha+)\Pi_r(u_s^r, \alpha) + p(u_s^r, u_s^r; \beta)(E[x|r, S^c(u_s^r, \alpha)] + \lambda(u_s^r, \alpha) E_A[x|r, S^c(u_s^r, \alpha)]). \tag{21}
\]

Divide each side of (19) by \( \alpha \) to get

\[
p(u_s, u_s^r; \beta) \left( E[x|s] - \frac{u_s}{\alpha} \right) = p(u_s^r, u_s^r; \beta)E[x|r, S^c(u_s^r, \alpha)] \tag{22}
\]

Substituting (22) into (21) and comparing it with (20), we obtain that (18) holds if and only if

\[
p_2(u_s, u_s; \beta)u_s'(\alpha+)\Pi_s(u_s, \alpha) + p(u_s, u_s; \beta)\frac{u_s}{\alpha} < p_2(u_s^r, u_s^r; \beta)u_s'(\alpha+)\Pi_r(u_s^r, \alpha) + p(u_s^r, u_s^r; \beta)\lambda(u_s^r, \alpha) E_A[x|r, S(u_s^r, \alpha)] \tag{23}
\]
To further simplify (23), note that
\[
\lambda(u^r_s, \alpha) = \frac{p_1(u^r_s, u_s; \beta)}{p(u^r_s, u_s; \beta)} \Pi_r(u^r_s, \alpha)
\]
\[
\Pi_s(u_s, \alpha) = \frac{p(u^r_s, u_s; \beta)}{p(u^r_s, u_s; \beta)} \Pi_r(u^r_s, \alpha)
\]
\[
p(u_s, u_s, \alpha) = \frac{p(u^r_s, u_s; \beta)\Pi_r(u^r_s, \alpha)}{\Pi_s(u_s, \alpha)} = p(u^r_s, u_s; \beta)\frac{p_1(u_s, u_s; \beta)}{p(u^r_s, u_s; \beta)} \Pi_r(u^r_s, \alpha)
\]
\[
u^r_s = E_A[x|r, S(u^r_s, \alpha)]
\]
The first equation uses the FOC \( u^r_s \) satisfies. The second equation is (19). The third equation uses (19) and the FOC \( u^r_s \) satisfies. The fourth is because the IR binds under the optimal wage scheme. Substituting these expressions to (23) and dividing each term by \( p(u^r_s, u_s; \beta)\Pi_r(u^r_s, \alpha) \), we obtain that (18) holds if and only if
\[
\frac{p_2(u_s, u_s; \beta)}{p(u_s, u_s; \beta)} u_s'(\alpha+) + \frac{p_1(u_s, u_s; \beta)}{p(u^r_s, u_s; \beta)} u_s^r < \frac{p_2(u^r_s, u_s; \beta)}{p(u^r_s, u_s; \beta)} u_s'(\alpha+) + \frac{p_1(u^r_s, u_s; \beta)}{p(u^r_s, u_s; \beta)} u_s^r.
\]
For \( u_s > 0 \), the above inequality follows from \( \partial \ln p/\partial u_1 \) being decreasing in \( u_1 \), (11), Remark 5.1, and that \( u^r_s > u_s \). For \( u_s = 0 \), the above inequality follows since \( \frac{\partial u_s}{\partial \alpha} = \frac{w_s}{\alpha} = 0 \). A similar argument holds for the left derivatives with respect to \( \alpha \).

Finally, consider the case where \( \alpha = \beta \) at the \( (\beta, \alpha) \) such that (19) holds. Then \( u_s(\beta, \alpha) \) is not differentiable, so the envelope theorem cannot be directly applied. However, a similar argument as in the proof of Lemma 3.2 can be used to show that single-crossing still holds at this point. Thus \( \beta_s \) is decreasing.

To show that \( \beta_r \) is decreasing, we want to show that as \( \alpha \) increases, deviations to safe contracts from a risky equilibrium become relatively less profitable. It suffices to show that whenever
\[
E\Pi_r(\beta, \alpha) = E\Pi^s_r(\beta, \alpha),
\]
we have
\[
\frac{\partial E\Pi_r}{\partial \alpha} > \frac{\partial E\Pi^s_r}{\partial \alpha}.
\]
An identical calculation as above shows that (24) is true if and only if
\[
\frac{p_2(u^s_r, u_s; \beta)}{p(u^s_r, u_s; \beta)} u^r_s(\alpha+) + \frac{p_1(u^s_r, u_s; \beta)}{p(u^r_s, u_s; \beta)} u^r_s > \frac{p_2(u_s, u_s; \beta)}{p(u^s_r, u_s; \beta)} u_s'(\alpha+) + \frac{p_1(u_s, u_s; \beta)}{p(u^s_r, u_s; \beta)} u_s^r.
\]
which follows from $\partial \ln \pi / \partial u_1$ being decreasing in $u_1$, (12), Remark 5.1, and that $u^*_r < u_r$.

Finally we are left to show that $\lim_{\alpha \to \infty} \beta_s(\alpha) = 0$. To start, observe that

$$\Pi_s(\alpha u^*, \alpha) = \Pi_r(\alpha u^*, \alpha).$$

where $u^*$ is the reservation utility that makes a principal indifferent between implementing $a = s$ or $a = r$ when there is no scaling: we can cancel out $\alpha$ on both sides of the equation and recover the same equation that defines $u^*$. Consequently, after scaling the principal is better off implementing the risky action whenever she has to offer the agent an indirect utility higher than $\alpha u^*$.

It then follows from (13) and the fact that $u^* < E[x | s]$ that for any $\beta > 0$, $u_s(\beta, \alpha) > \alpha u^*$ when $\alpha$ is sufficiently large. Therefore deviations to risky actions will eventually become profitable. This implies that $\lim_{\alpha \to \infty} \beta_s(\alpha) < \beta$ for all $\beta > 0$, and the claim follows. \hfill $\Box$

**Proof of Theorem 3.** Using the same method as in Proposition 3.1 we can show that $\Pi_r(\cdot, L)$ is concave for each $L$. Since $\Pi_r(E[x | s], L) > 0$ and $\Pi_s(E[x | s], L) = 0$, it suffices to prove the proposition for $u = 0$.\footnote{By concavity of $\Pi_r(\cdot, L)$ and the Mean Value Theorem argument in Proposition 3.1, $\Pi_r(u, L) > \Pi_s(u, L)$ for all $u \geq 0$.} Therefore we suppress $u$ and simply write $\Pi_r(L)$.

By the standard Lagrangian argument in Appendix A.1., the optimal risky contract is given by

$$w(x) = \begin{cases} 
-L, & x \in S^c(L) \\
x, & x \in S(L) 
\end{cases}$$

for a set $S(L)$ of the form $S(L) = A_{\lambda(L)} \cup (C_{\lambda(L)} \cap [0, q(L)])$ for some number $q(L)$. In particular, $S(L), S^c(L)$ are Borel sets.

Observe also that $\lambda(L)$ increases to $M$ as $L \to \infty$ and that $S(L)$ approaches the entire interval $[0, 1]$ in the sense that the measure of $S^c(L)$ goes to zero.

The profit under the optimal contract is given by

$$\Pi_r(L) = E[L + x | r, S^c(L)]$$

By concavity of $\Pi_r(\cdot, L)$ and the Mean Value Theorem argument in Proposition 3.1, $\Pi_r(u, L) > \Pi_s(u, L)$ for all $u \geq 0$. \hfill 44
s.t. $E_A[x|r, S(L)] = E_A[L|r, S^c(L)]$. In particular, $\Pi_r(L)$ is non-decreasing in $L$.

Substituting the agent’s IR constraint, we have

\[ \Pi_r(L) = \frac{P(S^c(L)|r)}{P_A(S^c(L)|r)} E_A[x|r, S(L)] + E[x|r, S^c(L)] \]

\[ = \int_{S^c(L)} \frac{f(x|r)}{f_A(x|r)} f_A(x|r)dx \frac{E_A[x|r, S(L)] + E[x|r, S^c(L)]}{P_A(S^c(L)|r)} \]

where $P_A(S^c(L)|r) = E_A[1|r, S^c(L)] = \int_{S^c(L)} f_A(x|r)dx$.

Fix an $\epsilon > 0$ so that $E_A[x|r] - \epsilon > E[x|s]$. Choose $L_1$ s.t. whenever $L > L_1$,

\[ E_A[x|r, S(L)] > E_A[x|r] - \epsilon \]

Choose $L_2$ such that $\lambda(L_2) > 1$. Then whenever $L > \max\{L_1, L_2\}$,

\[ \Pi_r(L) > E_A[x|r] - \epsilon > E[x|s] \]

Therefore, whenever $L$ is sufficiently large, any equilibrium must prescribe the risky action.

\[ \square \]

**Proof of Theorem 4.** Denote the principal’s optimal profit for implementing the risky action with indirect utility $u$ by $\Pi_r(u, L)$. Denote the relevant value functions by $E\Pi_s(u, L)$, $E\Pi^r_s(u, L)$, $E\Pi^r(u, L)$, $E\Pi_r(u, L)$ like before.

For each $L$, the proof of the existence of the cutoffs $\beta_s(L)$ and $\beta_r(L)$ follows from identical steps as in Theorem 1.

We argue that $\beta_s(\cdot)$ must decrease in $L$. This is because the payoff from the risky action deviation is higher at every level of $\beta$ than before, while $E\Pi_s(\beta, L)$ is the same for all $L \geq 0$ because $L$ does not affect the principal’s payoff from implementing the safe action. Therefore, for each $\beta$, the deviation to a risky contract from a safe equilibrium becomes more profitable as $L$ increases.

We next show that $\beta_r(\cdot)$ will decrease. For each $L$, let $u_r(\beta, L)$ be the symmetric equilibrium of the auxiliary game $p(u_i, u_{-i}; \beta)\Pi_r(u_i, L)$. Note that $u_r(\beta, \cdot)$ is increasing in $L$ if
\[ u_r(\beta, 0) > 0. \] To see this, note that the FOC for the symmetric equilibrium is

\[
\frac{p_1(u, u; \beta)}{p(u, u; \beta)} = \frac{\lambda(u, L)}{\Pi_r(u, L)}
\]

When \( L \) increases, the left-hand side is unchanged, but the right-hand side has a larger denominator and a smaller numerator. Thus the FOC will be satisfied at a higher level of \( u \).

Now we show that at every level of \( \beta \), the safe action deviation will be relatively worse than before. Suppose that at some \((\beta_r, L_r)\) the two functions intersect. Let \( u_r \) denote \( u_r(\beta_r, L_r) \) and \( u^*_r \) denote \( u^*_r(\beta_r, L_r) \).

Then by a similar argument as in Lemma A.5, we can show that the right derivatives with respect to \( L \) equal

\[
E\Pi'_r(\beta_r, L_r+) = p_2(u_r, u_r; \beta_r)u'_r(L_r+)\Pi_r(u_r, L_r) + p(u_r, u_r; \beta_r)\Pi'_r(u_r, L_r+)
\]

\[
(E\Pi^*_r)'(\beta_r, L_r+) = p_2(u^*_r, u_r; \beta_r)u'_r(L_r+)\Pi_s(u^*_r, L_r) + p(u^*_r, u_r; \beta_r)\Pi'_s(u^*_r, L_r+)
\]

We note that \( P'_r(u^*_r, L_r+) = 0 \) since the safe action is not affected by \( L \), and \( \frac{\partial \Pi(u_r, L_r)}{\partial L_r} > 0 \) since the risky action’s profit increases in \( L \). Then, since \( E\Pi_r(\beta_r, L_r) = p(u_r, u_r; \beta_r)\Pi_r(u_r, L_r) = p(u^*_r, u_r; \beta_r)\Pi_s(u^*_r, L_r) = E\Pi^*_r(\beta_r, L_r) \), after substitution and cancellation we obtain

\[
E\Pi'_r(\beta_r, L_r+) > (E\Pi^*_r)'(\beta_r, L_r+)
\]

\[
\iff \frac{p_2(u_r(\beta_r, L_r), u_r(\beta_r, L_r); \beta_r)}{p(u_r(\beta_r, L_r), u_r(\beta_r, L_r); \beta_r)} > \frac{p_2(u^*_r(\beta_r, L_r), u_r(\beta_r, L_r); \beta_r)}{p(u^*_r(\beta_r, L_r), u_r(\beta_r, L_r); \beta_r)}
\]

This holds from our assumptions on \( p \) since \( u^*_r(\beta_r, L_r) < u_r(\beta_r, L_r) \). A similar argument shows the inequality also holds for the left derivatives. Therefore whenever these two expected profit functions cross, the \( E\Pi_r(\beta_r, L_r) \) must cross from below.

Because of this single crossing from below, we know that higher levels of \( L \) will make the safe action relatively worse. Thus \( \beta_r(\cdot) \) is decreasing in \( L \).

\[ \square \]

**B Asymmetric Pure-strategy NEs Do Not Exist**

In this section we show that there is no asymmetric pure-strategy Nash equilibrium in the competing principals game with full information.
To do so, we establish two lemmas.

**Lemma B.1.** There is no equilibrium \((w_1, a_1), (w_2, a_2)\) such that \(a_1 = a_2\) but \(E_A[w_1|a_1] \neq E_A[w_2|a_2]\).

**Proof.** Assume to the contrary that \(a_1 = a_2 = a\) but \(u_1 = E_A[w_1|a] < u_2 = E_A[w_2|a]\). Since equilibrium is a mutual best response, \(u_1, u_2\) satisfy the FOCs

\[
\frac{p_1(u_1, u_2; \beta)}{p(u_1, u_2; \beta)} \leq \frac{-\Pi'_s(u_1)}{\Pi_s(u_1)}, u_1 \geq 0, \text{ with complementary slackness.}
\]

\[
\frac{p_1(u_2, u_1; \beta)}{p(u_2, u_1; \beta)} = \frac{-\Pi'_r(u_2)}{\Pi_r(u_2)}
\]

By Assumption 1, \(p_1/p\) is non-increasing in the first argument and non-decreasing in the second argument. Therefore we have

\[
\frac{p_1(u_1, u_2; \beta)}{p(u_1, u_2; \beta)} \geq \frac{p_1(u_2, u_1; \beta)}{p(u_2, u_1; \beta)}.
\]

Since \(-\Pi'_a(u)/\Pi_a(u)\) is increasing in \(u\) for either \(a = s\) or \(a = r\), we have

\[
\frac{-\Pi'_s(u_1)}{\Pi_s(u_1)} < \frac{-\Pi'_r(u_2)}{\Pi_r(u_2)},
\]

which is a contradiction.

Therefore, if there is an asymmetric equilibrium, the two principals must choose different \(a\)’s. To show that this case is also impossible, we need another lemma.

**Lemma B.2.** Let

\[
V_s(u_2) = \max_{u \in [0, E_A[x|s]]} p(u, u_2; \beta)\Pi_s(u)
\]

\[
V_r(u_2) = \max_{u \in [0, E_A[x|r]]} p(u, u_2; \beta)\Pi_r(u)
\]

where \(u_2 \geq 0\). Then if \(V_s, V_r\) ever cross each other, it must be that \(V_r\) crosses \(V_s\) from below.

**Proof.** Let \(u_s(u_2)\) be the maximizer of \(V_s(u_2)\) and \(u_r(u_2)\) be the maximizer of \(V_r(u_2)\). Then they satisfy the first-order condition

\[
\frac{p_1(u, u_2; \beta)}{p(u, u_2; \beta)} \leq \frac{-\Pi'_a(u)}{\Pi_a(u)}, u \geq 0, \text{ with complementary slackness}
\]
for \( a = s, r \) respectively. The proof of Lemma A.6 shows that \( 1/\Pi_s(u) > -\Pi_r'(u)/\Pi_r(u) \) for all \( u \geq 0 \). Therefore, \( u_s(u_2) < u_r(u_2) \) for all \( u_2 \geq 0 \).

Suppose \( V_r, V_s \) cross at \( \hat{u} \). Then

\[
p(u_s(\hat{u}), \hat{u}; \beta)\Pi_s(u_s(\hat{u})) = p(u_r(\hat{u}), \hat{u}; \beta)\Pi_r(u_r(\hat{u}))
\]

(25)

By the Envelope Theorem,

\[
V'_s(\hat{u}) = p_2(u_s(\hat{u}), \hat{u}; \beta)\Pi_s(u_s(\hat{u}))
\]

\[
V'_r(\hat{u}) = p_2(u_r(\hat{u}), \hat{u}; \beta)\Pi_r(u_r(\hat{u}))
\]

Substituting (25), we then have

\[
V'_r(\hat{u}) > V'_s(\hat{u})
\]

\[
\iff \frac{p_2(u_s(\hat{u}), \hat{u}; \beta)}{p(u_s(\hat{u}), \hat{u}; \beta)} < \frac{p_2(u_r(\hat{u}), \hat{u}; \beta)}{p(u_r(\hat{u}), \hat{u}; \beta)}
\]

By Assumption 3, the latter inequality is true. This completes the proof.

Now we are in a position to show the non-existence result.

**Proposition B.1.** There exists no asymmetric pure-strategy Nash equilibrium in the competing-principals game with full information.

**Proof.** By Lemma B.1, if there is an asymmetric equilibrium, it must be of the form \((w_1, s), (w_2, r)\). Let \( u_s = E_A[w_1(x)|s] \) and \( u_r = E_A[w_2(x)|r] \). By Proposition 3.1,

\[
u_s \leq u^* \leq u_r
\]

(26)

**Claim:** At least an inequality in (26) must be strict.

To see this, note that if \( u_r = u_s = u^* \), then because \((w_1, s)\) is a best response,

\[
p_1(u^*, u^*; \beta)\Pi_s(u^*) + p\Pi'_s(u^*) = 0
\]

(27)

Since \( \Pi'_r(u^*) > \Pi'_s(u^*) = -1 \) (See Figure 1), (27) leads to

\[
p_1(u^*, u^*; \beta)\Pi_r(u^*) + p\Pi'_r(u^*) > 0
\]
which contradicts the fact that \((w_2, r)\) is a best response. This proves the claim. Accordingly, 
\(u_s < u_r\).

Since \((w_1, s), (w_2, r)\) is an equilibrium, each principal has no incentive to deviate to a different action \(a\). That is to say,

\[
\begin{align*}
V_r(u_s) &\geq V_s(u_s) \\
V_s(u_r) &\geq V_r(u_r).
\end{align*}
\] (28)

Since \(u_s < u_r\), (28) implies that \(V_r\) crosses \(V_s\) from above, which contradicts Lemma B.2.

\(\square\)
References


