Robust Security Design

Seokwoo Lee*† Uday Rajan‡

April, 2018

Abstract

We consider the optimal contract between an entrepreneur and investors in a single-period model when both parties have limited liability, are risk-neutral toward cash flow risk, and are ambiguity-averse. Ambiguity aversion is modeled by multiplier preferences for robustness toward model uncertainty, as in Hansen and Sargent (2001). Efficient ambiguity-sharing implies that the first-best contract consists of either convertible debt or levered equity. As is customary, in the second-best contract, moral hazard is alleviated by giving more cash to investors in low cash flow states. Under many settings in our model, the optimal security has an equity-like component in high cash flow states, providing a contrast to the results in Innes (1990). Finally, we derive the optimal contracts when investors can acquire information that reduces the amount of ambiguity over cash flows and the two parties can renegotiate the initial contract. We find that the optimal initial contract is risky debt, and this is renegotiated to either convertible debt or levered equity.

Keyword: Robustness, Multiplier Preference, Ambiguity Aversion, Moral Hazard, Renegotiation, Security Design, Convertibility

*We are grateful to Archishman Chakraborty, Simon Gervais, William Mann (discussant), Alejandro Rivera (discussant), Francesco Sangiorgi, Martin Szydlowski (discussant), and conference participants at the Minnesota Corporate Finance Conference and FTG London meetings, EFA and Northern Finance Association for helpful comments.

†George Mason University; slee65@gmu.edu

‡Stephen M. Ross School of Business, University of Michigan; urajan@umich.edu

Uncertainty is one of the fundamental facts of life. It is as ineradicable from business decisions as from those in any other field.

—Frank H. Knight (1921), Part III, Chapter XII.

1 Introduction

Startup firms face uncertain futures. To be successful, a startup firm must provide consumers with a new product or service. It is difficult to predict many of the factors that affect the cash flows to the new firm. These factors may be external (the degree to which consumers will like the new product, the response by rival firms currently in the market, the possibility of future disruptive technological change) or internal (the ability to execute on a strategic plan and to manage growth) to the startup firm. The firm therefore faces uncertainty in the sense of Knight (1921)—an inability to quantify the probabilities over different future outcomes.[1] Indeed, according to Knight, profit is a reward to an entrepreneur for bearing uncertainty.

As the Ellsberg (1961) paradox demonstrates, individuals are averse to uncertainty, preferring gambles with known probabilities to those with unknown probabilities. Gilboa and Schmeidler (1989) resolve the Ellsberg paradox by postulating a multiple priors model. A subject does not have enough information to form a prior belief; rather, she has in mind a set of prior distributions and believes that any one of them may be the true prior. Further, she is averse to this ambiguity, and evaluates a gamble according to the minimal expected utility over all priors in this set. Hansen and Sargent (2001) extend the maxmin expected utility notion of Gilboa and Schmeidler by adding on a penalty function for evaluating gambles according to different distributions, based on the distance of a distribution from some reference measure. The interpretation is that the agent understands that her reference model may be misspecified, and wishes to make a decision that is robust to an error in specifying the model.

In this paper, we consider the implications of ambiguity-aversion on the part of both an entrepreneur and investors for security design. We build upon the model of Innes (1990). The entrepreneur has a project for which he needs to raise external financing. After the investment, the entrepreneur takes a costly action that affects the distribution of future cash flows from the project. Both parties are risk-neutral; that is, for both parties, the value of a gamble with known probabilities is equal to the expected cash flow from the gamble. In addition, both have limited liability. However, both are ambiguity-averse. We adopt the approach to ambiguity aversion of (Hansen and Sargent, 2001). The investors and the entrepreneur in our model are concerned about model misspecification, and are averse to this prospect.

In the Innes (1990) model, if incentive compatibility binds, the optimal security gives all cash to investors in low cash flow states and all cash to the entrepreneur in high states. If the security is required to have non-decreasing payments, the optimal security is a debt contract, leaving all

[1] In contrast, outcomes with known probabilities are termed "risky" rather than "uncertain."

cash to the investors in low states and a fixed payment to them in high states. The model is both elegant and simple.

However, most venture capital contracts do not conform to its stark predictions. Indeed, Gompers and Lerner (2001) define venture capital in terms of its "focus on equity or equity-linked investments" in private high-growth firms. Kaplan and Strömberg (2003) report that, in *all* of the 213 deals in their sample, venture capitalists retain substantial cash flow rights in high states of the world. The most common form of security in their sample is convertible preferred stock, with the investor having the option to convert to common stock in case of an exit (such as an IPO or an acquisition by another firm). The few deals that do not include any convertible security include common stock as one of the securities issued to investors. There is, therefore, a substantial gap between the results of the Innes model and actual securities used in venture capital transactions.

As Schmidt (2003) and Hellmann (2006) show, double moral hazard (i.e., moral hazard on the part of both entrepreneur and investors) leads to the use of convertible securities in venture capital settings. We provide a different explanation for the existence of convertible features—ambiguity aversion on both sides.

Our main insight is that ambiguity-aversion generates gains to trade from ambiguity-sharing, which generally necessitates the presence of equity-like features in the optimal security. We begin by considering the first-best case, in which the entrepreneur's effort is directly contractible. In this case, the optimal security involves sharing cash flow proportionally between the investors and the entrepreneur in high cash flow states. Thus, the optimal security directly has an equity component. Depending on how cash flows are split up in low states of the world, it is interpretable as either convertible debt or levered equity, with unlevered equity arising in a knife-edge case.

When effort is not contractible, the contract must induce an incentive compatible action from the entrepreneur. Compared to the security in the first-best contract, this requires the investors to obtain more cash in low cash flow states. We demonstrate the conditions that must be satisfied by the optimal contract (which consists of a security issued to investors and the level of effort to be undertaken by the entrepreneur), and generate a number of numerical examples to understand the features of the security and the comparative statics of the problem.

We find that the division of cash flow in high cash flow states depends on the degree of ambiguity aversion of both investors and entrepreneur. First, suppose that the entrepreneur has high ambiguity-aversion. The security is not linear in this case due to the binding incentive compatibility constraint, but has an equity-like feature to the extent that the payment increases in the cash flow from the project.[2] As investors become less ambiguity-averse, they have a greater preference (relative to the entrepreneur) for receiving cash in high states, which increases the slope of the optimal security. Next, suppose the entrepreneur has low ambiguity-aversion. Relative to the

[2] Convertible securities in the venture capital setting often have complicated participation features that can create non-linearities in their payments; see Kaplan and Strömberg (2003).

investors, the entrepreneur prefers to receive cash in high states. In addition solving the incentive problem requires withholding cash from the entrepreneur in low states. These two factors reinforce each other, so the security held by the investors can have a non-monotonic payment, with regions in which its payment decreases as the project cash flow increases.

Finally, we require the investors and the entrepreneur to both hold claims whose payments are non-decreasing in the project cash flow. Then, in the latter case in which the entrepreneur is not too ambiguity-averse, the optimal security is debt (recovering the Innes result) if investors are extremely ambiguity-averse. If investors too are not very ambiguity-averse, the optimal security is convertible debt, with conversion feasible only sufficiently high states.

Overall, we find that debt emerges as the optimal security only if the entrepreneur has a low degree of ambiguity aversion and the investors are significantly more ambiguity-averse. In all other settings, the security offers some payment to the investors in high cash flow states, which is interpretable as an equity component. Further, when the entrepreneur is sufficiently ambiguity-averse, the security offers non-decreasing payments without requiring that feature to be imposed, and has a strong resemblance to equity.

We adopt the Hansen and Sargent (2001) approach to modeling ambiguity aversion, specifically using what they term "multiplier preferences," to capture the notion that the agent cares about robustness toward model misspecification. This approach embeds the Gilboa and Schmeidler (1989) maxmin approach to ambiguity aversion in a tractable setting in which an additional parameter describes the degree of aversion the agent has toward model uncertainty. An alternative would be to model the degree of ambiguity aversion using the smooth ambiguity aversion approach of Klibanoff et al. (2005); we find the Hansen and Sargent approach more tractable in our setting. Maccheroni et al. (2006) provide an axiomatization of variational preferences (a broader set), and Strzalecki (2011) extends the axiomatization to multiplier preferences.

Hansen and Sargent (2001) show that multiplier preferences are closely related to constraint preferences. The multiplier (i.e., the parameter which captures the degree of ambiguity aversion) in the former case is interpretable as the shadow cost of the information constraint in the latter. This shadow cost, in turn, must depend on the amount of information available to the decision-maker. Information about the cash flows of a small firm improves over time, which will then imply decreasing informational costs, or decreasing multipliers. This motivates us to consider renegotiation of the financial contract when ambiguity aversion reduces over time.

Our renegotiation model is built on Hermalin and Katz (1991) and Dewatripont et al. (2003). Suppose that after the entrepreneur chooses effort, but before the firm's cash flow is realized, the investor observes the chosen level of effort. This effort remains non-verifiable. At this point, the amount of uncertainty faced by the contracting parties decreases, which is also interpretable as a reduction in their degrees of ambiguity aversion. Here, the crucial aspect of renegotiation is that it separates the incentive and insurance problems. Initially, the entrepreneur offers a debt contract, which makes the entrepreneur the residual claimant, and so provides strong incentives to provide effort. After the effort is observed, the parties renegotiate this debt contract to an efficient ambiguity-sharing contract. Importantly, in our framework, this contract is *linear*: Renegotiation eliminates non-linear components in the second-best contract, because incentive compatibility is no longer issue at the renegotiation. As a result, convertible debt emerges optimally with renegotiation. We demonstrate via an example that the initial debt contract may be renegotiated to levered equity, especially if the investor is relatively less ambiguity-averse than the entrepreneur at the renegotiation stage.

In our model, the entrepreneur and the investors are both ambiguity-averse, which yields gains to trade from ambiguity-sharing. In the venture capital context, venture capitalists are often thought to have a better knowledge of the industry whereas an entrepreneur may know more about their own firm's technology. Thus, it is reasonable to think of both investors and entrepreneur as facing model uncertainty.[3] Miao and Rivera (2016) and Szydlowski (2012), study robust contracts in continuous time, assuming that the principal alone (but not the agent) faces uncertainty. In Miao and Rivera (2016), the principal does not know the output distribution chosen by the agent, but the agent does. They determine the optimal dynamic contract and exhibit an implementation featuring cash, debt, and equity. Szydlowski (2012) models a cost of effort for the agent that changes over time. The agent naturally knows his own effort cost, but the principal is uncertain about it. Ambiguity leads to excessive compensation following a high performance, and under-compensation following a low performance.

Bewley (1989) builds a theory of innovation and entrepreneurship based on uncertainty aversion. In his model, entrepreneurship is undertaken by individual investors with low levels of uncertainty aversion. In our model, the entrepreneur and investors are distinct entities. Arguably, in the venture capital arena, uncertainty aversion is low (relative to the population) among both investors and founders.

In a related paper, Andrey Malenko (20018) investigate optimal security design by a perfectly informed (and therefore ambiguity-neutral) issuer when the investor faces Knightian uncertainty about the firm's cash flow distribution. In their framework, the equilibrium security depends on both the issuer's private information and the degree of uncertainty faced by the investor. They find that debt, equity, or call options can emerge as the optimal security. As the issuer is ambiguity-neutral and the investor ambiguity-averse, ambiguity-sharing in their paper takes on a different form as compared to our model with smooth preferences toward uncertainty. A more important distinction is the presence of private information for the issuer in their model. In contrast, we work with a model in which information is symmetric but there can be moral hazard on the part of the

[3]There is, of course, a distinction between whether an agent faces model uncertainty and whether the agent is averse to that uncertainty. The latter is a preference characteristic, and an agent may well be ambiguity-neutral despite facing a high degree of uncertainty.

entrepreneur.

In many settings in our context, the optimal security has equity-like features. As mentioned earlier, Schmidt (2003) and Hellmann (2006) explain these features in venture capital contracts on the basis of double moral hazard. Convertible debt also emerges if the entrepreneur is risk-averse and contracts can be renegotiated (Dewatripont et al. (2003)) or if the investor is risk-averse and the entrepreneur can engage in risk-shifting (Ozerturk (2008)). Stein (1992) takes a different approach to the use of convertible debt by large firms. In his model, convertible debt is a form of backdoor equity financing, to avoid the usual adverse selection discount to equity. Ortner and Schmalz (2016) consider an optimistic entrepreneur issuing a security to an investor who is more pessimistic, and show that convertible debt can emerge when the project has an embedded expansion option. We note that ambiguity aversion may be thought of as a micro-foundation for belief disagreement. To the extent that agents have different degrees of ambiguity aversion, they are effectively evaluating outcomes under different probability distributions.

Our paper adds to the recent literature on the ambiguity aversion in corporate finance settings. For example, in the model of Dicks and Fulghieri (2015), ambiguity-aversion leads to endogenous disagreement between firm insiders and external shareholders, thus creating a motive for governance. Relatedly, Garlappi et al. (2015) show that, in settings such as corporate boards, the group in the aggregate can act like an ambiguity-averse decision-maker. Ambiguity-aversion also explains innovation and merger waves, by generating a strategic complementarity in investment in innovative projects (Dicks and Fulghieri (2016)).

The rest of this paper is organized as follows. We provide a brief introduction to the Hansen and Sargent (2001) multiplier preferences approach in Section 2. The model is introduced in Section 3, and the first-best contract is described. The solution to the full contracting problem is exhibited in Section 4, with and without a monotonicity requirement on security payments. Section 5 addresses the distinctions from risk-aversion. The solutions to the optimal contracting with ex-post renegotiation is presented in Section 6. Section 7 concludes.

2 Multiplier Preferences and Ambiguity Aversion

In this section, we briefly review multiplier preferences, which were introduced by Hansen and Sargent (2001) to capture model uncertainty; that is, the notion that a decision-maker does not know the true probability distribution of events, and is averse to model misspecification.

Consider a set of states (events) X. We define a payoff profile $r : X \mapsto Z$, where Z is a set of consequences, and define Bernoulli utility function $u : Z \mapsto \mathbb{R}$. Let Σ denote a sigma-algebra of events in X. Let $\Delta(X)$ denote the set of all countably-additive probability measures on X. Given a measure q, with a slight abuse of notation, let $\Delta(q)$ denote the set of probability measures equivalent to q.

Then, given a probability measure q, an expected utility maximizer evaluates a payoff profile r

according to the criterion $U(r) = \int_X u(r(x))\, dq(x)$. The expected utility maximizer prefers payoff profile r_1 to r_2 if and only if $U(r_1) \geq U(r_2)$.

Now, consider a decision-maker who has a reference probability measure q, but is uncertain about the true measure. Here, q may be thought of as the decision-maker's "best guess" about the true probabilities over events. Given any other measure $p \in \Delta(q)$, the relative entropy $R(p||q)$ is defined by

$$R(p||q) = \begin{cases} \int_X \left(\ln \frac{p(x)}{q(x)}\right) dp(x) & \text{if } p \in \Delta(q) \\ \infty & \text{otherwise} \end{cases} \tag{1}$$

The relative entropy $R(\cdot||q)$ (also called the Kullback-Leibler divergence between p and q) provides a distance metric between p and q. It is non-negative, and equal to zero if and only if $p = q$ (see Dupuis and Ellis (1997), Lemma 1.4.1). Moreover, it is convex in p (see Dupuis and Ellis (1997), Lemma 1.4.3).

According to the multiplier preferences introduced by Hansen and Sargent (2001), when faced with a payoff profile with a reference measure q, the decision-maker allows for the notion that his reference measure may be incorrect, and therefore allows himself to evaluate the payoff profile according to some other measure p that is close to q. Probability measures far from q are given greater weight in the decision, and so are considered more costly to choose. Specifically, the decision-maker evaluates a payoff profile r with reference measure q according to

$$V(r) = \min_{p \in \Delta(q)} \int_X u(r(x)) dp(x) + \theta R(p||q), \tag{2}$$

where $\theta > 0$. A payoff profile r_1 is preferred to r_2 if and only if $V(r_1) \geq V(r_2)$, and the decision-maker's goal is to maximize V.

Here, θ is inversely related to the degree of ambiguity aversion on the part of the decision-maker (see Maccheroni et al. (2006), Corollary 21). Specifically, it captures the extent of the decision-maker's aversion to the risk that the model (or reference measure q) has been misspecified. Intuitively, one can think of the decision-maker choosing a distribution p according to which to evaluate the lottery q. As θ becomes large, the penalty for choosing a distribution far from the reference distribution q increases, which naturally leads to a distribution closer to q being chosen. That is, as θ becomes large, the decision-maker is less concerned with model misspecification (as they believe that the true measure is close to the reference measure), or is less ambiguity-averse.

In the limit as $\theta \to \infty$, the probability distribution p that minimizes the right-hand-side of equation (2) must equal q, so we have $V(r) = U(r)$ for a given payoff profile r. That is, the decision criterion reduces to the usual notion of maximizing expected utility. Conversely, as $\theta \to 0$, the decision-maker becomes infinitely ambiguity averse.

From Dupuis and Ellis (1997), Proposition 1.4.2 (see also Strzalecki (2011), Section 3.3),

$$\min_{p\in\Delta(q)} \int_X u(r(x))dp(x) + \theta R(p||q) \quad = \quad -\theta \ln\left(\int_X e^{-\frac{u(r(x))}{\theta}} dq(x)\right). \tag{3}$$

Therefore, a decision-maker maximizing the LHS of equation (3) may equivalently be modeled as maximizing the RHS, so that we can write

$$V(r) \quad = \quad -\theta \ln\left(\int_X e^{-\frac{u(r(x))}{\theta}} dq(x)\right). \tag{4}$$

Going forward, for the rest of the paper, we will assume that all parties have ambiguity-averse preferences represented as in equation (4).

3 Security Design Problem

We build upon the model of Innes (1990). A penniless entrepreneur has a project that requires an investment I at date 0. The investment I must therefore be raised from external investors. The project generates a cash flow $x \in X = [0, \overline{x}]$ at date 1, which is then shared between the investors and the entrepreneur. By assumption, the cash flow x is non-negative. Let $r(x)$ denote the amount given to the investors, and $w(x) = x - r(x)$ the amount retained by the entrepreneur. Both entrepreneur and investors have limited liability, so $0 \leq r(x) \leq x$ for all x. The function $r(\cdot)$ is naturally interpretable as a financial security, so that the choice of r is a security design problem.

After the investment is undertaken, the entrepreneur takes an action, or equivalently provides an effort, $a \geq 0$, which incurs a utility cost $\psi(a)$. The cost is strictly increasing and strictly convex in a, so that $\psi'(a) > 0$ and $\psi''(a) > 0$. In addition, we assume that $\psi(0) = 0$. Investors and entrepreneurs agree on the effect that the cost has on the reference measure induced over the cash flows by the action a. In particular, they believe that action a likely leads to a distribution $F(x \mid a)$ over cash flows at date 1, with associated density $f(x \mid a)$. We assume that $F(x \mid a)$ has full support over X and has no mass points.

We assume that $f(x \mid a)$ satisfies the Monotone Likelihood Ratio Property (MLRP). Denote $f_a(x \mid a) = \frac{\partial f(x|a)}{\partial a}$. Then, $\frac{\partial}{\partial a}\left(\frac{f_a(x|a)}{f(x|a)}\right) > 0$.

The entrepreneur and the investors are both neutral toward cash flow risk, so that $u(x) = x$ for both parties. However, both are ambiguity-averse in the sense of being averse to risk of model misspecification. Investors value risky cash flows according to equation (4), with ambiguity-aversion parameter θ_I. That is, the value to investors of a security $r(x)$ when the action is a is given by

$$V_I(r, a) \quad = \quad -\theta_I \ln\left(\int_X e^{-\frac{r(x)}{\theta_I}} f(x \mid a)dx\right) \tag{5}$$

Similarly, the entrepreneur evaluates risky cash flows according to equation (4), with ambiguity-aversion parameter θ_E. In addition, the entrepreneur privately bears the cost of the action, $\psi(a)$. Therefore, given an action a and a security $r(x)$ offered to investors, the entrepreneur's value for the contract is

$$V_E(r,a) \quad = \quad -\theta_E \ln\left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx\right) - \psi(a). \tag{6}$$

Without loss of generality, we set the discount rate between date 0 and date 1 to zero. The investment I has no uncertainty associated with it. It is straightforward to see that $-\theta_I \ln\left(\int_X e^{-\frac{I}{\theta_I}} f(x \mid a)dx\right) = I$, so that we can write the investors' individual rationality, or IR, constraint as:

$$-\theta_I \ln\left(\int_X e^{-\frac{r(x)}{\theta_I}} f(x \mid a)dx\right) \quad \geq \quad I. \tag{7}$$

Because the action a is taken after the investment has been made, it cannot be committed to by the entrepreneur. Instead, as is usual, a must be incentive compatible. The relevant incentive compatibility (IC) condition for the entrepreneur is

$$a \quad = \quad \arg\max_{\tilde{a}} -\theta_E \ln\left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid \tilde{a})dx\right) - \psi(\tilde{a}). \tag{8}$$

For now, we assume the first-order approach is valid (later we specify a sufficient condition for this). We therefore replace the IC condition in equation (8) with the corresponding first-order condition

$$-\theta_E \frac{\int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x \mid a)dx}{\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx} - \psi'(a) \quad = \quad 0. \tag{9}$$

The complete contracting problem may therefore be stated as:

$$\text{[Problem P1]} \quad \max_{r(x),a} \quad -\theta_E \ln\left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx\right) - \psi(a) \tag{10}$$

$$\text{subject to:} \qquad \text{(IR)} \quad -\theta_I \ln\left(\int_X e^{-\frac{r(x)}{\theta_I}} f(x \mid a)dx\right) \geq I \tag{11}$$

$$\text{(IC)} \quad -\theta_E \frac{\int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x|a)dx}{\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x|a)dx} - \psi'(a) = 0. \tag{12}$$

$$\text{(LL)} \qquad 0 \leq r(x) \leq x \text{ for all } x. \tag{13}$$

Here, (LL) represents the limited liability constraints on the security. We refer to the pair (r,a) as a contract, and to r as a security.

We assume that a feasible solution exists to this problem. Essentially, that requires that, given the conditional density over cash flows $f(x \mid a)$, the effort cost function $\psi(a)$, and the set of feasible cash flows X, the required investment level I is sufficiently low.

We first transform the maximization problem P1 into an equivalent minimization problem P2 that does not require the use of natural logs. The benefit is that when we determine the first-order conditions in r and a, the corresponding derivatives have a simpler form.

$$\text{[Problem P2]} \qquad \min_{r(x),a} \quad e^{\frac{\psi(a)}{\theta_E}} \left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a) dx \right) \tag{14}$$

$$\text{subject to: (IR2)} \qquad \int_X e^{-\frac{r(x)}{\theta_I}} f(x \mid a) dx \; \leq \; e^{-\frac{I}{\theta_I}} \tag{15}$$

$$\text{(IC2)} \quad \int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x \mid a) dx + \frac{\psi'(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a) dx \; = \; 0 \tag{16}$$

$$\text{(LL)} \qquad 0 \leq r(x) \leq x \text{ for all } x. \tag{17}$$

Lemma 1. *Problems P1 and P2 have the same set of solutions.*

All proofs are contained in the Appendix.

Observe that the transformed functions for the entrepreneur and investor in (14) and (15) are identical to utility functions that represent constant absolute risk aversion (CARA) preferences. We comment further on this in Section 5.

Before we exhibit the optimal contract in our model, we briefly review the results from Innes (1990). Our model is identical to the Innes model except for the feature of ambiguity aversion on the part of investors and entrepreneur. In the limit, as $\theta_I \to \infty$ and $\theta_E \to \infty$, investors and entrepreneurs become ambiguity-neutral in our model, so that in the limiting case the model reduces exactly to the Innes model.

Three benchmark results from Innes (1990) are of interest to us: (1) In the Innes model, in the first-best outcome security design is irrelevant, as both investors and entrepreneur are risk-neutral. That is, as long as the effort is at the first-best level and the investors' IR constraint holds, any division of project cash flows between the two parties is optimal. (2) In the second-best problem, if the entrepreneur's incentive compatibility (IC) condition binds, the optimal security provides all cash flows to investors in low states, and all cash flows to the entrepreneur in high states.[4] (3) If the security held by investors must provide payments to them that are weakly monotone in the cash flow x, the optimal security is debt.

3.1 First-best Problem

In the first-best problem, incentive compatibility is not an issue, or, put another way, we can think of the action as being directly contractible. The contract can specify an effort level a, and a security r that specifies cash flows to investors, contingent on the cash flows of the project, if the entrepreneur in fact chooses action a. As effort is contractible, if the entrepreneur chooses

[4]In the second-best problem, incentive compatibility may or may not bind, depending on how high I is relative to the distribution over cash flows at date 1, given the optimal effort level.

any action $\tilde{a} \neq a$, investors can give the entrepreneur zero cash and retain the entire output x for themselves.

Let λ denote the shadow price for the investors' IR constraint in equation (15). Further, for each x, let $\underline{\gamma}_x$ denote the shadow price on the constraint $r(x) \geq 0$ and $\overline{\gamma}_x$ the shadow price for the constraint $r(x) \leq x$. Then, the Lagrangian for the first-best problem may be written as:

$$\begin{aligned} \mathcal{L}_f(r, a, \lambda) &= e^{\frac{\psi(a)}{\theta_E}} \left(\int_X e^{-\frac{x - r(x)}{\theta_E}} f(x \mid a) dx \right) + \lambda \left[\int_X e^{-\frac{r(x)}{\theta_I}} f(x \mid a) dx - e^{-\frac{I}{\theta_I}} \right] \\ &\quad + \int_X \left[-\underline{\gamma}_x r(x) + \overline{\gamma}_x (r(x) - x) \right] dx \end{aligned} \tag{18}$$

Let a_f denote the optimal effort level in the first-best problem, r_f the optimal security, and λ_f the shadow price of the investors' IR constraint given the first-best contract. Finally, given $y \in R$, let $y^+ = \max\{0, y\}$.

We show that the solution to the first-best contract produces a security that is piecewise-linear in the project cash flow x. The entrepreneur and investors face two sources of unexpected outcomes in this problem: cash flow risk and model uncertainty. Cash flow risk is represented by the reference density $f(x \mid a)$, and both parties are neutral toward it. Model uncertainty implies that the true cash flow distribution may be different from the reference distribution, and investors are averse to it. The latter creates a motive for ambiguity-sharing that leads to an outcome in which high cash flows are shared between investors and entrepreneur. Depending on the ambiguity aversion of each side, low cash flows may be given entirely to the investors or entirely to the entrepreneurs.

Proposition 1. *In any solution to the first-best problem,*

(i) The investors' IR constraint binds.

(ii) The optimal security satisfies

$$r_f(x) = \min \left\{ x, \left(\frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I \theta_E}{\theta_I + \theta_E} \left(\ln \frac{\lambda_f \theta_E}{\theta_I} - \ln e^{\frac{\psi(a_f)}{\theta_E}} \right) \right)^+ \right\}. \tag{19}$$

Suppose that for some value of x, we have a strictly interior solution for $r_f(x)$; that is, $r_f(x) \in (0, x)$. Equation (19) says that in this case, $r_f(x) = \frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I \theta_E}{\theta_I + \theta_E} \left(\ln \frac{\lambda_f \theta_E}{\theta_I} - \ln e^{\frac{\psi(a_f)}{\theta_E}} \right)$. Observe that the term inside the parentheses does not depend on x, so that r_f is linear in x.[5] The linear term $\frac{\theta_I}{\theta_I + \theta_E} x$ reflects optimal ambiguity-sharing between investors and entrepreneur. The linearity of the term follows from the exponential form of the expressions in problem P2, the transformed

[5]Also, of course, $\ln e^{\frac{\psi(a_f)}{\theta_E}} = \frac{\psi(a_f)}{\theta_E}$. We state the contract using the expression $\ln e^{\frac{\psi(a_f)}{\theta_E}}$ to facilitate comparison with the second-best contract in the next section.

problem.[6] It follows that the payoff on the security is weakly increasing in x, and is overall piecewise linear in x.

The intuition behind the first term $\frac{\theta_I}{\theta_I+\theta_E}\,x$ is as follows. Recall that θ_I is inversely related to the degree of ambiguity aversion expressed by the investors, and likewise θ_E is inversely related to the ambiguity aversion of the entrepreneur. The more ambiguity-averse an agent is (i.e., the lower θ is), the further (and so the more pessimistic) the distribution under which they evaluate the cash flows is, compared to the reference measure $f(x \mid a)$. A pessimistic agent places greater weight on low cash flow outcomes, and so prefers to receive cash in those low states. Conversely, a less ambiguity-averse agent is relatively optimistic, and so prefers to receive cash in high cash flow states.

Now, keeping the optimal effort a_f fixed, as θ_I increases, keeping θ_E fixed, the investors become less ambiguity-averse relative to the entrepreneur. Optimal ambiguity-sharing thus entails that the slope of the investors' share of the cash flow increases, so that investors get relatively more cash in the high states. Of course, as θ_I increases, the gains to ambiguity-sharing between the investors and entrepreneur also change, which has a feedback effect on the optimal effort in the first-best problem.

In addition to the investors' IR constraint, equation (15), and the form of the security in equation (19) (both mentioned in the statement of Proposition 1) the optimal action a_f satisfies the first-order condition $\frac{\partial \mathcal{L}_f}{\partial a} = 0$, or

$$e^{\frac{\psi(a)}{\theta_E}}\left(\frac{\psi'(a)}{\theta_E}\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx + \int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x \mid a)dx\right) + \lambda \int_X e^{-\frac{r(x)}{\theta_I}} f_a(x \mid a)dx = 0. \quad (20)$$

These three equations, (15), (19), and (20) can be used to solve for a_f, λ_f, and r_f.

Note that we have assumed in our formulation that investors and entrepreneur have the same reference density in mind $f(x \mid a)$. Suppose, for example, investors evaluated cash flows based on a reference density $f_I(x \mid a) \neq f(x \mid a)$. In that case, the relative optimism or pessimism of each party will depend both on the coefficient θ and the reference measure used to evaluate cash flows. As a result, the optimal security will depend at any x on both $f_I(x \mid a)$ and $f(x \mid a)$.

3.2 Example 1

To illustrate the properties and comparative statics of the first-best contract, we consider the following numeric example. Let $X = [0, 1]$, and let the action set be $A = [0, 1]$. Set $f(x \mid a) = 1 + a(2x - 1)$, so that $f_a(x \mid a) = 2x - 1$ and $f_{aa}(x \mid a) = 0$. Note that $\frac{f_a(x|a)}{f(x|a)} = \frac{1}{a+\frac{1}{2x-1}}$, which is clearly increasing in x, so that MLRP is satisfied. Let $\psi(a) = \frac{1}{2}a^2$, so that $\psi'(a) = a$ and $\psi''(a) = 1$. Finally, let $I = 0.3$.

[6] Recall that, as shown by Wilson (1968), optimal risk-sharing with exponential utilities entails a linear sharing rule.

When both parties are ambiguity-neutral, i.e., in the case of the Innes (1990) model, the first-best effort is found by solving the first-order condition $\int_X x f_a(x \mid a)dx = \psi'(a)$, which in this example yields $a_N^* = \frac{1}{6}$ (the subscript $_N$ denotes that both parties are ambiguity-neutral). Any division of the cash flows such that in expectation the investors obtain I and the entrepreneur obtains $\int_X x f(x \mid a) - I$ is optimal.

With ambiguity-aversion there are three possibilities for the optimal security issued to the investors in the first-best case in this example. We illustrate these three cases by keeping θ_E fixed at 1, and varying θ_I. In each case, the security issued in the first-best case includes a substantial equity component.

1. Convertible debt.

   This security emerges if θ_I is sufficiently low, relative to θ_E. As investors are pessimistic relative to entrepreneurs, in the low cash flow states all cash is given to the investors. Their financial claim therefore resembles debt in the low states. In the high cash flow states, the motive for ambiguity-sharing kicks in, and cash flows are divided between investors and entrepreneur using the linear sharing rule mentioned above. That is, once the cash flow exceeds a threshold, both investors and entrepreneur own equity in the project. Putting the two pieces together, the security is convertible debt with the conversion threshold set equal to the face value of the debt.

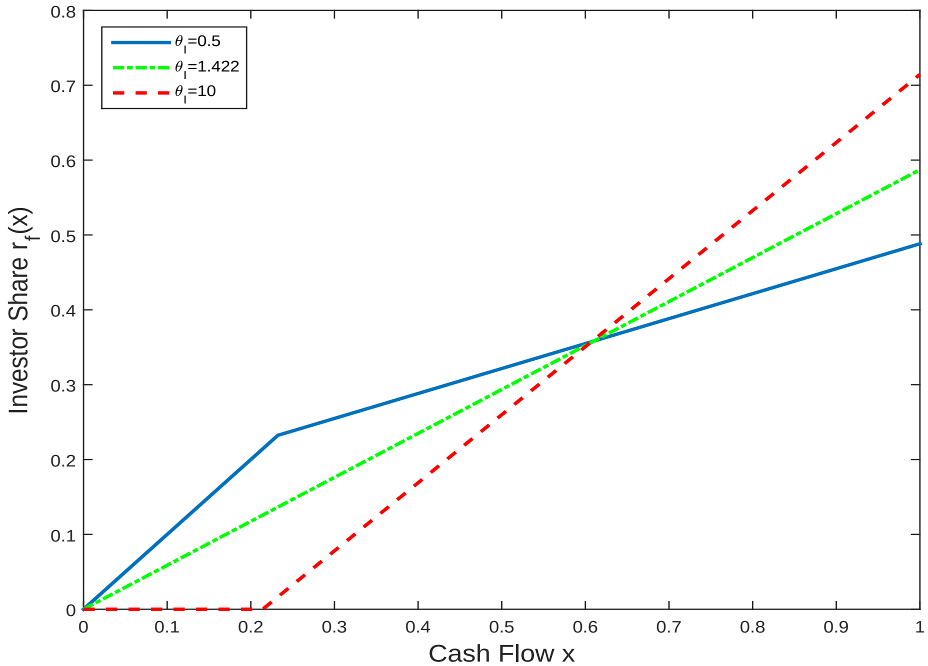
2. Levered Equity.

   This security emerges if θ_I is sufficiently high, relative to θ_E. Here, the entrepreneur is pessimistic relative to investors, and obtains all cash in the low states. Once cash flow is sufficiently high, we are back to the case in which ambiguity-sharing adds value, with both parties holding equity claims. The security can therefore be characterized as levered equity, with the entrepreneur holding priority over cash flows in low states.

3. Unlevered Equity.

   This is a knife-edge case that emerges at a specific value of θ_I; in the example, at θ_I approximately equal to 1.422.

In each of the three cases, the equity fraction the investor obtains in the region in which cash flows are shared is given by $\frac{\theta_I}{\theta_I+\theta_E}$. The entrepreneur has a financial claim that is the mirror image of that issued to the investor. In the case that the entrepreneur has convertible debt, of course, his financial claim can equivalently be interpreted as a salary (subject to the firm having the cash to pay the salary) plus a stock bonus. We illustrate the different financial securities that emerge as θ_I varies in Figure 1.



This figure illustrates the securities issued to the investor as θ_I varies. We set $f(x \mid a) = 1 + a(2x - 1)$, $\psi(a) = \frac{1}{2}a^2$, $I = 0.3$, and $\theta_E = 1$.

Figure 1: Securities Issued to Investor in First-Best Contract

The optimal effort in the first-best contract falls as θ_I increases, which is intuitive. An increase in θ_I implies that the entrepreneur receives more cash in the low cash flow states. Therefore, the incentive to provide effort to reach the higher cash flow states is lower. Note that in this case, the "total surplus" from the first-best contract depends on the preferences of both investors and entrepreneur (because that determines the gains to ambiguity-sharing). Therefore, it depends also on the level of investment, which affects how the gains from ambiguity-sharing are divided. As a result, as shown in Table 1 below, the effort level for some parameter values may be greater than the level when both parties are ambiguity neutral, $\frac{1}{6}$.

4 Second-best Problem

We now turn to the second-best problem. Recall that in this case effort is not directly contractible, but rather must be chosen so as to be incentive compatible for the entrepreneur. Taking into account the entrepreneur's IC constraint in equation (16), the Lagrangian for problem P2 may be written as:

$$\begin{aligned}\mathcal{L}(r,a,\lambda) &= e^{\frac{\psi(a)}{\theta_E}}\left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x\mid a)dx\right)+\lambda\left[\int_X e^{-\frac{r(x)}{\theta_I}} f(x\mid a)dx - e^{-\frac{I}{\theta_I}}\right]\\ &\quad +\mu\left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x\mid a)dx+\frac{\psi'(a)}{\theta_E}\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x\mid a)dx\right)\\ &\quad +\int_X\left[-\underline{\gamma}_x r(x)+\overline{\gamma}_x(r(x)-x)\right]dx. \qquad (21)\end{aligned}$$

Here, μ is the shadow price on the entrepreneur's IC constraint, and, as before, λ is the shadow price on the investors' IR constraint, $\underline{\gamma}_x$ the shadow price on the constraint $r(x)\geq 0$, and $\overline{\gamma}_x$ the shadow price on the constraint $r(x)\leq x$.

As we show, the optimal security in the second-best problem entails a weak reduction in the cash flow paid to the investors in high states. When the IC condition binds, the reduction is strict. Essentially, relative to the first-best problem, more cash must be given to the entrepreneur in the high states to induce effort.

Denote with a^* superscript the value of a variable in a solution to the second-best problem.

Proposition 2. *In any solution to the second-best problem,*

(i) The investors' IR constraint binds.

(ii) The optimal security satisfies

$$r^*(x)=\min\left\{x,\left(\frac{\theta_I}{\theta_I+\theta_E}x+\frac{\theta_I\theta_E}{\theta_I+\theta_E}\left\{\ln\frac{\lambda^*\theta_E}{\theta_I}-\ln\left(e^{\frac{\psi(a^*)}{\theta_E}}+\mu^*\left(\frac{f_a(x\mid a^*)}{f(x\mid a^*)}+\frac{\psi'(a^*)}{\theta_E}\right)\right)\right\}\right)^+\right\}. \qquad (22)$$

For any value of x at which the cash flow is divided between investors and entrepreneur, equation (22) implies that $r^*(x)=\frac{\theta_I}{\theta_I+\theta_E}x+\frac{\theta_I\theta_E}{\theta_I+\theta_E}\left\{\ln\frac{\lambda^*\theta_E}{\theta_I}-\ln\left(e^{\frac{\psi(a^*)}{\theta_E}}+\mu^*\left(\frac{f_a(x|a^*)}{f(x|a^*)}+\frac{\psi'(a^*)}{\theta_E}\right)\right)\right\}$. As before, the term $\frac{\theta_I}{\theta_I+\theta_E}x$ is linear and increasing in x. However, the term $\frac{f_a(x|a^*)}{f(x|a^*)}$ is also increasing in x, so when $\mu>0$, the term in the curly parentheses is decreasing in x in some non-linear fashion.

As Innes (1990) shows, incentive compatibility may not bind in this case. In the Innes model, whether IC binds depends on the level of investment, I, relative to the density over cash flows, $f(x\mid a)$. When I is low, IC does not bind, so that the contract reverts to a first-best contract. However, when I is high, IC binds and $\mu>0$. A similar intuition goes through with ambiguity aversion. We have shown that the first-best contract in our setting is piecewise linear and has an equity component. For the rest of the paper, we concentrate on the case that the IC constraint binds.

Two implications emerge when $\mu>0$: First, the security held by the investors overall is no longer piecewise linear in x, and can have significant non-linear components. Therefore, the security is no longer directly interpretable in terms of equity, although it can have an equity-like component. Second, the payoff on the security need not be weakly increasing in project cash flow—in particular, there may exist ranges of cash flow such that the investors' payout is decreasing as x increases. We demonstrate this property in the context of a numerical example in the next section.

In the second-best case, the first-order condition in a is $\frac{\partial L}{\partial a} = 0$, which reduces to

$$\lambda \int_X e^{-\frac{r(x)}{\theta_I}} f_a(x \mid a)dx + \mu \left\{ \frac{\psi''(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx + \frac{\psi'(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x \mid a)dx \right.$$
$$\left. + \int_X e^{-\frac{x-r(x)}{\theta_E}} f_{aa}(x \mid a)dx \right\} = 0. \quad (23)$$

The four conditions represented by the above equation, the investors' IR condition (15), the entrepreneur's IC condition (16), and the equation for the optimal contract (22) can be used to solve for a^*, λ^*, μ^*, and r^* in the case that the IC condition binds, so that $\mu > 0$.

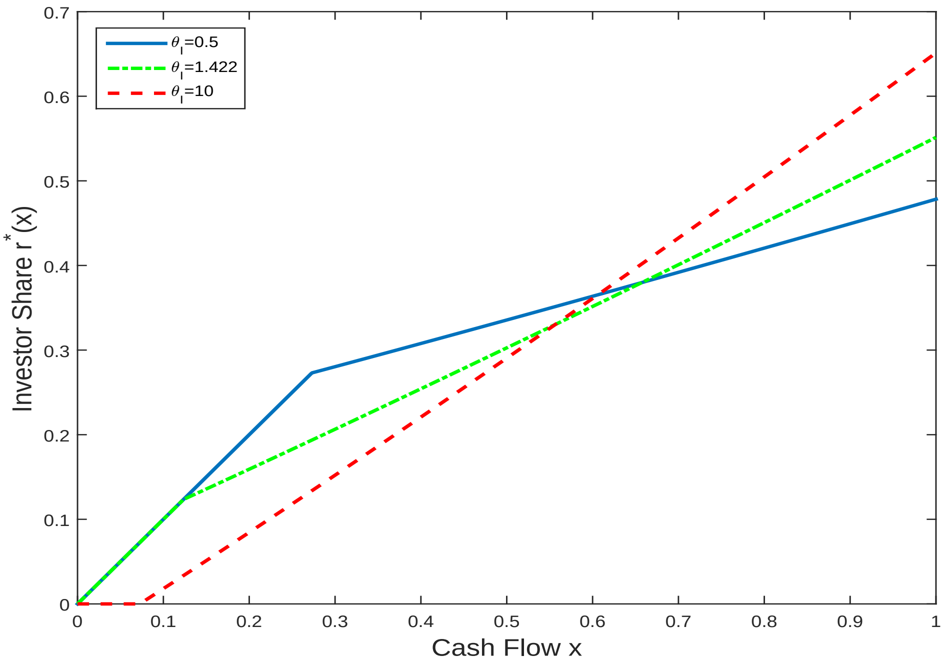
4.1 Example 2

We use the same parameters as in Example 1. We set $\theta_E = 1, I = 0.3, f(x \mid a) = 1 + a(2x - 1)$, and $\psi(a) = \frac{1}{2}a^2$. The first-best security for three different values of θ_I is exhibited in Figure 1. We illustrate the optimal security in the second-best setting in Figure 2. To illustrate the difference between the first and second-best contracts, we choose parameter values at which the IC constraint binds in the second-best problem.

In contrast to the security in the first-best case, the optimal security in the second-best case provides more cash flow to the investors in low states. This is true for all three levels of θ_I. In the figures, the contrast is greatest for the intermediate θ_I case (with $\theta_I = 1.422$), with the first-best contract entailing straight equity, but the security in the second-best contract resembling convertible debt. Note that the securities in Figure 2 are *not* piecewise linear—for high cash flows, there is a slight non-linearity in the security payoffs. Therefore, they cannot be thought of directly in terms of equity. Nevertheless, for these parameter values, the securities have a component that resembles equity to a large degree.

There is a natural tension in the problem between ambiguity-sharing and the need to provide incentives to the entrepreneur. That is, the usual trade-off between risk and incentives is resurrected by as a trade-off between uncertainty and incentives. On the one hand, if the entrepreneur were ambiguity-neutral, the moral hazard problem would entail giving the entrepreneur less cash in low states. On the other hand, optimal ambiguity-sharing involves the investor receiving less cash in low states and more cash in high states. The design of the security, in turn, feeds back into the moral hazard problem, and affects the optimal effort provided by the agent.

We report the optimal effort levels in the first- and second-best problems in our example in



This figure illustrates the securities issued to the investor as θ_I varies. We set $f(x \mid a) = 1 + a(2x - 1)$, $\psi(a) = \frac{1}{2}a^2$, $I = 0.3$, and $\theta_E = 1$.

Figure 2: Securities Issued to Investor in Second-Best Contract

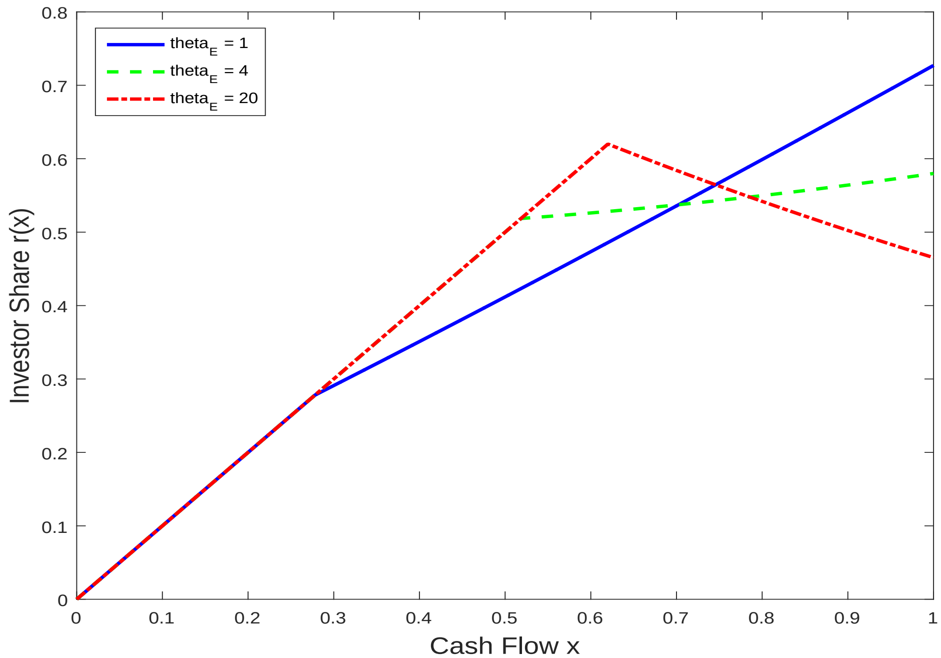
θ_I	0.5	1.422	10
First-best effort	0.176	0.168	0.165
Second-best effort	0.096	0.082	0.052

Table 1: Optimal Effort Levels

Table 1.

Next, we consider the comparative statics of the optimal security as θ_E changes. We use the same cost function, conditional cash flow density, and investment level as before. We set $\theta_I = 4$ and vary θ_E across three levels, 1, 4, and 20. The optimal security in each case is exhibited in Figure 3.

The broad intuition is as follows. The optimal security balances out the need for ambiguity-sharing, which entails giving more cash to the less ambiguity-averse party in high states, with the need to provide incentives to the entrepreneur, which entails giving more (often all) cash to the investor in low states. As θ_E increases, the entrepreneur becomes less ambiguity-averse, and so starts to get paid in lower states. Further, the slope of the optimal security falls, reflecting the fact that the entrepreneur obtains a greater proportion of the cash in high states.



This figure illustrates how the optimal security changes as θ_E changes. Throughout, we set $\theta_I = 4, I = 0.3, \psi(a) = \frac{1}{2}a^2$, and $f(x \mid a) = 1 + a(2x - 1)$.

Figure 3: Optimal Security as θ_E Changes

When the entrepreneur has sufficiently lower ambiguity aversion than the entrepreneur, the optimal security can have a payout that over some range is decreasing in cash flow, a security that provides the entrepreneur with large amounts of cash in the high cash flow states provides the best incentives, because the entrepreneur is relatively confident in the reference probability measure $f(x \mid a)$. In our example, when $\theta_E = 20$, the security payoffs decrease in the project cash flow when x exceeds approximately 0.6.

Finally, we note that, although it is not immediate from the figure, the security payoffs are non-linear in x in the region in which the cash flow is being shared between investors and entrepreneur.

4.2 Second-Best Contract with Monotone Security Payoffs

As shown in Proposition 1, the security contained in the first-best contract has a payoff that is weakly increasing in x. Further, there exists a threshold $\hat{x}$ such that the security is strictly increasing and linear in x for $x \geq \hat{x}$. Further, in many settings, the security issued in the second-best contract also has a payoff weakly increasing in x. For example, fixing a value of θ_I, if θ_E is sufficiently low, a security with increasing payoffs provides the entrepreneur with a hedge against model uncertainty.

Nevertheless, as shown in Figure 3, when θ_E is high relative to θ_I, a security with payoffs that decrease when x is high provides the entrepreneur with high-powered incentives. In such situations, we consider the implications of introducing another restriction on the security, that the payoffs must be non-decreasing in cash flow.

We say a security is monotone in the cash flow x if both the investor and the entrepreneur obtain a cash payment that is non-decreasing in x.

Definition 1. *A security $r(x)$ is monotone in the cash flow x if $r(x)$ and $w(x) = x - r(x)$ are each non-decreasing in x.*

Note that monotonicity implies that the security must be continuous in x, and also must be differentiable almost everywhere (i.e., except over a set of measure zero). We therefore operationalize the monotonicity condition by adding an extra condition (M) to problem P2, that $r'(x) \geq 0$ for almost all x. In addition, as $w(x)$ is non-decreasing, it must be that $r'(x) \leq 1$. The full contracting problem is then:

$$
\begin{array}{lll}
\text{[Problem P3]} & & \min_{r(x),a} \quad e^{\frac{\psi(a)}{\theta_E}} \left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a) dx \right) \\
\text{subject to:} & \text{(IR2)} & \int_X e^{-\frac{r(x)}{\theta_I}} f(x \mid a) dx \;\leq\; e^{-\frac{I}{\theta_I}} \\
& \text{(IC2)} & \int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x \mid a) dx + \frac{\psi'(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a) dx \;=\; 0 \\
& \text{(LL)} & 0 \leq r(x) \leq x \text{ for all } x. \\
& \text{(M)} & 0 \leq r'(x) \leq 1 \text{ for almost all } x.
\end{array}
$$

Given Proposition 2, denote

$$
\tilde{r}(x) = \frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I \theta_E}{\theta_I + \theta_E} \left\{ \ln \frac{\lambda \theta_E}{\theta_I} - \ln \left(e^{\frac{\psi(a)}{\theta_E}} + \mu \left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E} \right) \right) \right\}. \tag{24}
$$

Further, define $\hat{r}(x) = \min\{\max\{\tilde{r}(x), 0\}, x\}$, so that $\hat{r}(x)$ ensures that limited liability is satisfied for both entrepreneur and investors.

At any given x at which $\hat{r}(x)$ is differentiable, there are two possibilities: (i) $\hat{r}'(x) \geq 0$, in which case condition (M) is satisfied at that value of x, or (ii) $\hat{r}'(x) < 0$, in which case condition (M) is violated at that x. In the latter case, define $m(x) = \sup\{y \leq x \mid r'(y) \geq 0\}$. Finally, define

$$
r^M(x) \quad = \quad \begin{cases} \hat{r}(x) & \text{if } \hat{r}'(x) \geq 0 \\ \hat{r}(m(x)) & \text{otherwise.} \end{cases} \tag{25}
$$

Then, the optimal contract once condition (M) is added to the problem includes $\hat{r}^M$ as the security issued to the investors. Let (a^M, λ^M, μ^M) denote values at which problem P3 is solved, given that the security issued is r^M.

Proposition 3. *In any solution to problem P3,* *(i) The investors' IR constraint binds.*

(ii) The security issued to the investors is $r^M(x)$, evaluated at (a^M, λ^M, μ^M).

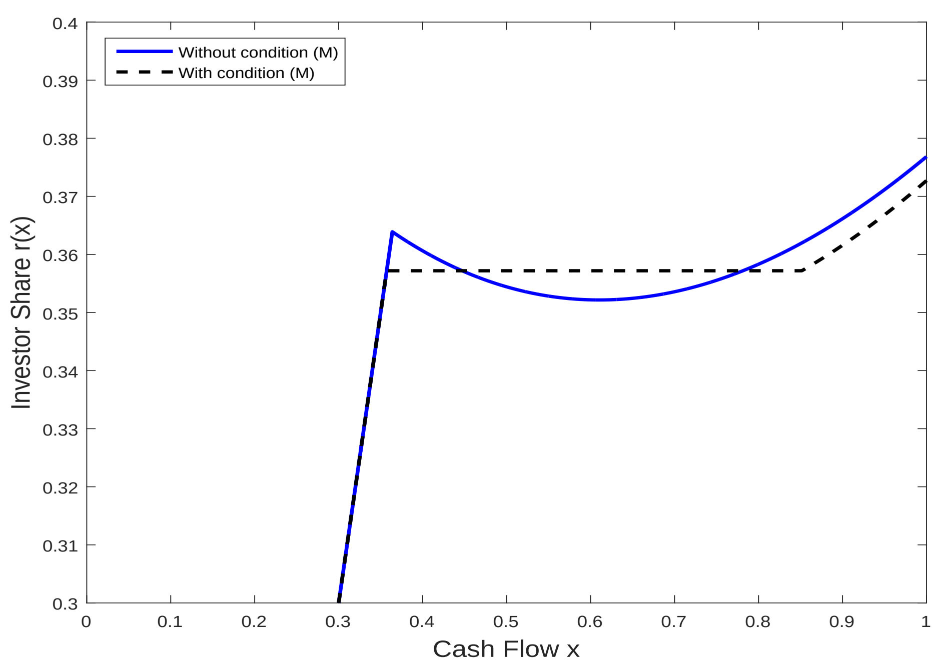
That is, the payoff of the security at any cash flow x is determined as in Proposition 3, and any decreasing portions are "flattened" so that the eventual payoff is weakly increasing. Of course, in doing so, the values of a, λ, and μ will all change.

The proof of the proposition uses optimal control techniques similar to those used in ironing in a standard mechanism design problem with unknown types.

To illustrate the application of Proposition 3, we revert to our running example. Let $\psi(a) = \frac{1}{2}a^2$ and $f(x \mid a) = 1 + a(2x - 1)$. Consider two examples of monotonic contracts.

First, set $\theta_I = 4$ and $\theta_E = 20$. In Figure 3, we exhibit the payoff on the optimal security when monotonicity is not imposed. As seen from the figure, the security has decreasing payoffs for x greater than approximately 0.6. If the security is required to be monotone, the optimal security is debt. It has a payoff $r(x) = x$ for $x \leq$ approximately 0.53, and $r(x) \approx 0.53$ for $x > 0.36$.

As a second example, set $\theta_I = 20$ and $\theta_E = 10.5$. We exhibit the second-best contract with and without condition (M) in Figure 4. To highlight the effects of condition (M), we change the scale of the Y-axis to display the relevant region of $r^M(x)$.



This figure illustrates the effect of requiring the security to have weakly increasing payoffs. We set $\theta_I = 20, \theta_E = 10.5, I = 0.3, \psi(a) = \frac{1}{2}a^2$, and $f(x \mid a) = 1 + a(2x - 1)$.

Figure 4: Optimal Security With and Without Monotonicity Requirement

Overall, then, we find that when the entrepreneur is ambiguity-averse (i.e., θ_E is low) and investors are mildly ambiguity-averse relative to the entrepreneur (i.e., θ_I is sufficiently high relative to θ_E), the optimal security has strictly increasing payoffs. When the entrepreneur is mildly ambiguity-averse (i.e., θ_E is high), the optimal security can have segments in which its payoff decreases in project cash flow. In this case, introducing the additional restriction that security payoffs are non-decreasing in x leads to either standard debt (when investors are ambiguity-averse, so θ_I is low) or a security that resembles convertible debt with a high strike price for the conversion option (when investors are only mildly ambiguity-averse, so that θ_I is high).

4.3 Validity of first-order approach

We now identify a sufficient condition for the first-order approach to incentive compatibility to be valid in our problem. As in equation (6), let V_E denote the entrepreneur's payoff. Let $w(x) = x - r(x)$; then $V_E = -\theta_E \ln\left(\int_X e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{w(x)}{\theta_E}} f(x \mid a)dx\right) - \psi(a) = -\theta_E \ln\left(e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{w(x)}{\theta_E}} f(x \mid a)dx\right)$. Denote $U = e^{-\frac{V_E}{\theta_E}} = e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{w(x)}{\theta_E}} f(x \mid a)dx$. Then, if action a maximizes V_E, it must minimize U.

A sufficient condition for U to have a local minimum in a is $\frac{\partial^2 U}{\partial a^2} > 0$. Differentiate U twice with respect to a. Then, at any point at which the entrepreneurs' IC condition is satisfied, we obtain $\frac{\partial^2 U}{\partial a^2} = e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{w(x)}{\theta_E}} f(x \mid a) \left[\frac{\psi''(a)}{\theta_E} + \frac{\psi'(a)}{\theta_E}\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{f_{aa}(x \mid a)}{f(x \mid a)}\right] dx$.

Because $e^{\frac{\psi(a)}{\theta_E}} > 0$, for $\frac{\partial^2 U}{\partial a^2} > 0$, it is sufficient for a local minimum that

$$\int_X e^{-\frac{w(x)}{\theta_E}} f(x \mid a)\left[\frac{\psi''(a)}{\theta_E} + \frac{\psi'(a)}{\theta_E}\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{f_{aa}(x \mid a)}{f(x \mid a)}\right]dx > 0 \tag{26}$$

If the equation is satisfied for all a, then U is strictly concave in a, so we have a global minimum.

Note that a sufficient condition for (26) to be satisfied at a given effort level a is that

$$\psi''(a) + \psi'(a)\frac{f_a(x \mid a)}{f(x \mid a)} + \theta_E \frac{f_{aa}(x \mid a)}{f(x \mid a)} \quad \geq \quad 0 \text{ for all } x. \tag{27}$$

Condition (26) can also be expressed in terms of the distribution function F. Applying integration by parts repeatedly to equation (26) and simplifying, an equivalent sufficient condition is

$$e^{-\frac{w(x)}{\theta_E}} \frac{\psi''(a)}{\theta_E} + \int_X \frac{w'(x)}{\theta_E} e^{-\frac{w(x)}{\theta_E}} F(x \mid a)\left(\frac{\psi''(a)}{\theta_E} + \frac{\psi'(a)}{\theta_E}\frac{F_a(x \mid a)}{F(x \mid a)} + \frac{F_{aa}(x \mid a)}{F(x \mid a)}\right)dx \quad > \quad 0 \tag{28}$$

This condition is satisfied at a given effort level a if $\psi''(a) + \psi'(a)\theta_E \frac{F_a(x|a)}{F(x|a)} + \theta_E \frac{F_{aa}(x|a)}{F(x|a)} > 0$ for all x.

In our numerical examples, we set $f(x \mid a) = 1 + a(2x - 1)$, so that $f_a(x \mid a) = 2x - 1$ and

$f_{aa}(x \mid a) = 0$. Therefore, the minimum value of $\frac{f_a(x|a)}{f(x|a)}$ is $-\frac{1}{1-a}$, attained when $x = 0$. Further, $\psi(a) = \frac{1}{2}a^2$, so that $\psi'(a) = a$ and $\psi''(a) = 1$. Therefore, in the examples,

$$\psi''(a) + \psi'(a)\frac{f_a(x \mid a)}{f(x \mid a)} + \theta_E \frac{f_{aa}(x \mid a)}{f(x \mid a)} = 1 + \frac{a}{a + \frac{1}{2x-1}}. \tag{29}$$

The minimum value of the RHS is $1 - \frac{a}{1-a}$, attained when $x = 0$. Therefore, if $a \leq \frac{1}{2}$, condition (27) is satisfied for all x. The entrepreneur's payoff is therefore concave for $a \in [0, 0.5]$. In the examples, the values of effort we find are considerably less than 0.5, so we have identified a minimum over the range $[0, 0.5]$. Further, the marginal cost of effort at $a = 0.5$ is sufficiently high that higher values of a are not optimal for the entrepreneur.

5 Discussion: Interpretation of θ

As noted before, the framework we are using, with risk-neutrality and ambiguity-aversion, recovers a functional form similar to that of constant absolute risk aversion (CARA) utility to represent the preferences of the entrepreneur and the investor. In this section, we emphasize that the interpretation of θ is quite different under ambiguity aversion, as compared to under risk aversion.

First, we note that the exponential functional form arises in the Hansen and Sargent (2001) framework because, under multiplier preferences, the relative entropy $R(p||q)$ is chosen to measure the distance between the measure p used to evaluate a stochastic payoff and the reference measure q. Maccheroni et al. (2006) show that multiplier preferences are a special case of a larger set of preferences, variational preferences. Variational preferences are represented more generally as

$$V(r) = \min_{p \in \Delta(q)} \int_X u(r(x))dp(x) + \theta\, c(p, q),$$

where $c : \Delta(q) \to [0, \infty]$ is a convex function on $\Delta(q)$. This general representation does not admit an equivalent exponential form if $c(\cdot, q) \neq R(\cdot||q)$.

Hansen and Sargent (2001), in Chapter 6, show that multiplier preferences are closely related to constraint preferences, which may be written as:

$$\hat{V}(r) = \min_{p \in \Delta(q)} \int_X u(r(x))dp(x), \qquad \text{subject to } R(p||q) \leq \eta \tag{30}$$

In this variant, the parameter θ is interpretable as the shadow cost on the information or relative entropy constraint. The right-hand side of the constraint, η, is then inversely related to the multiplier θ.[7] For example, as $\eta \to 0$ (or, as noted on page 6, when $\theta \to \infty$), in the limit maximizing $\hat{V}(r)$ is equivalent to maximizing expected utility.

Now, consider a firm evolving through time. As time increases and more information about the

[7] See Appendix A.

firm's future prospects becomes available, it seems natural to think that the amount of uncertainty about the firm goes down. In the constraint preference model, this would correspond to η decreasing, even though the underlying preferences over ambiguity remain unchanged. In our formulation with variational preferences, this would translate to the parameter θ increasing over time. If θ is instead interpreted as the coefficient of absolute risk aversion, it would remain unchanged if preferences are invariant over time. We build upon the idea of time-varying θ in the next section, where we consider renegotiation between the entrepreneur and the investor when new information arrives at some time beyond when the initial contract is signed.

6 Renegotiation with Time-varying θ

We now consider optimal contracts in a setting in which θ increases over time. In particular, we assume that at date 0.5, the investors obtain some information about the cash flow at time 1, which results in an increase in their ambiguity-aversion parameter. Specifically, their ambiguity aversion parameter increases from θ_{I0} to θ_{I1}. For simplicity, we assume that the entrepreneur's ambiguity aversion parameter, θ_E, remains unchanged, as does the reference probability measure q used by both parties.

The increase in θ_I may then generate gains to trade between the entrepreneur and the investor, which allows room for renegotiating the optimal contract. Following Hermalin and Katz (1991) and Dewatripont et al. (2003), to keep the renegotiation problem tractable, we also assume that at this stage the investor observes the effort that the entrepreneur provided at date 0. However, this effort is non-verifiable and hence non-contractible. Of course, at this stage the effort is sunk, and hence there is no longer any need to provide incentives to the agent. Thus, the focus of the renegotiation is on finding an efficient way to share ambiguity and risk between the entrepreneur and the investor.

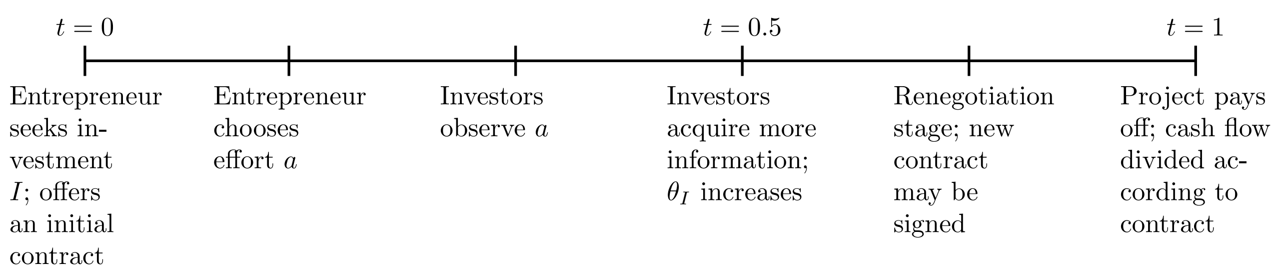


Figure 5: Sequence of events with renegotiation

We assume that the entrepreneur has all bargaining power at the renegotiation stage. That is, the entrepreneur makes a take-it-or-leave-it offer to the investors at this stage. The investors (who at this point have observed the effort incurred at time 0) will accept the renegotiation offer if and only if their utility from the new contract is at least as high as their utility from the old contract.

If they reject the offer, the initial contract prevails.

6.1 Optimal securities with ex-post renegotiation

Our approach in this section parallels that in Dewatripont et al. (2003). Recall that, in order to provide incentives to the entrepreneur, the second-best contract turns out to be non-linear (Proposition 2). In particular, although it can have an equity-like component, it also contains non-linear components. Our main result in this section is that renegotiation eliminates such non-linear components. Effort is no an longer issue at the renegotiation stage. As a result, as we show in Proposition 4 below, the parties renegotiate the initial contract to an efficient ambiguity-sharing contract that is indeed linear.

Let $r_0(\cdot)$ and $r_n(\cdot)$ respectively be the initial and the renegotiated contracts that the entrepreneur offers to the investors. The entrepreneur's choice of $r_n(\cdot)$ depends both on the initial contract $r_0(\cdot)$ and on the effort a. Let $\hat{u}(\cdot)$ be inversely related to investors' utility given the old contract $r_0(\cdot)$, effort a (which is known to the investor before renegotiation occurs), and their ambiguity aversion parameter. Specifically, define

$$\hat{u}(r_0(\cdot), a; \theta_{I1}) := \int_X e^{-\frac{r_0(x)}{\theta_{I1}}} f(x \mid a) dx \tag{31}$$

The utility level implied by $\hat{u}$ serves as the reservation utility of the investors at the renegotiation stage.

The intuition for the optimal security is as follows. All else equal (in particular, holding fixed the action a), the entrepreneur will be better off at the renegotiation stage if $\hat{u}$ is maximal across all feasible contracts offered at date 0. That is, the lower the utility level investors can be held to, the higher the value to the entrepreneur. In particular, (i) fixing a, the security at date 0 that maximizes $\hat{u}$ is risky debt, and (ii) once $r_0(\cdot)$ and a have been chosen, at the renegotiation stage the new security $r_n(\cdot)$ solves the first-best problem given a and $\hat{u}$. Given these two points, the security design problem at date 0 reduces to choosing the face value of risky debt such that the induced effort by the entrepreneur maximizes the value of the contract to the entrepreneur.

To formalize this intuition, let $C[0, x]$ be the set of continuous functions defined on the domain $[0, \overline{x}]$. The set of feasible contracts that satisfies limited liability and monotonicity is defined as:

$$\mathcal{C} := \left\{ r(\cdot) \in C[0, \overline{x}] \mid 0 \leq r(x) \leq x \text{ for all } x,\ r'(x) \in [0,1] \text{ for almost all } x \right\}. \tag{32}$$

The transformed value functional for the entrepreneur at the renegotiation-stage is then:

$$T(\hat{u}, a) \;:=\; \min_{r_n(\cdot)\in\mathcal{C}} \left\{ e^{\frac{\psi(a)}{\theta_E}} \left(\int_X e^{-\frac{x-r_n(x)}{\theta_E}} f(x \mid a) dx \right) \quad \text{subject to} \right.$$

$$\left. \int_X e^{-\frac{r_n(x)}{\theta_{I1}}} f(x \mid a) dx \;\leq\; \hat{u}(r_0(\cdot), a; \theta_{I1}) \right\}. \tag{33}$$

At date 0, the entrepreneur chooses the optimal initial contract $r_0(\cdot)$ and the optimal effort a to minimize the transformed value functional $T(\hat{u}, a)$. Investors must, of course, obtain at least their reservation utility. After renegotiation, investors will be held down to the utility implied by $\hat{u}$. Therefore, at the minimum, individual rationality of investors must require that $\hat{u}(r_0(\cdot), a;\, \theta_{I1}) \leq e^{-\frac{I}{\theta_I}}$, as in the definition of $T(\cdot)$ above. We go one step further, and require that the security $r_0(\cdot)$ must satisfy the investors' individual rationality constraint at the initial ambiguity aversion parameter θ_{I0} and chosen effort level a. That is, we require that $\hat{u}(r_0(\cdot), a;\, \theta_{I0}) \leq e^{-\frac{I}{\theta_{I0}}}$, as in condition (15). The justification for this assumption is that although we model renegotiation as costless, in practice there is always a chance that renegotiation is unsuccessful. We can view the requirement that $r_0(\cdot)$ satisfy investor individual rationality as equivalent to one in which there is a vanishing probability that renegotiation is unsuccessful, and that every final contract must satisfy the investors' IR constraint.

As $T(\cdot)$ is inversely related to the entrepreneur's value from the contract, the entrepreneur's problem at date 0 can be stated as:

$$\min_{r_0(\cdot)\in\mathcal{C},\, a} \; T(\hat{u},\, a), \tag{34}$$

where $\mathcal{C}$ as defined in equation (32) requires contracts to satisfy limited liability and monotonicity.

We start by showing that indeed the entrepreneur's transformed value functional is decreasing, and in addition is convex, in $\hat{u}$.

Lemma 2. *Keeping the effort a fixed, the transformed value functional $T(\cdot)$ is convex and decreasing in $\hat{u}(\cdot)$.*

Now, suppose we fix the initial contract r_0 and increase the effort level a. Observe that, with limited liability and the lowest cash flow being $0 < I$, any feasible contract must be risky. Thus, as a increases, MLRP implies $\hat{u}(r_0, a)$ in turn decreases. That is, effort by the entrepreneur has the effect of increasing the investors' reservation utility at the renegotiation stage. To minimize this effect, relative to the first-best problem (in which effort is directly contractible), there is under-provision of effort given the optimal contract.

Intuitively, as $T(\cdot)$ is decreasing in $\hat{u}$, the optimal security should minimize the rate at which $\hat{u}$ falls as the effort increases; that is, should minimize $|\frac{\partial \hat{u}}{\partial a}|$. We show in the proof of the next

Proposition that the optimal security at date 0 is risky debt, which is of course a piecewise linear contract. At the renegotiation stage, the security is converted to an optimal ambiguity-sharing security that is also piecewise-linear, and hence has a substantial equity component. Thus, at both the initial and renegotiation stages, the optimal contracts are piecewise-linear.

Proposition 4. *Suppose the initial contract too must satisfy limited liability. Then, the optimal initial security is risky debt with a suitably chosen face value D^*, so that $r_0^*(x) = \min\{x, D^*\}$. Further,*

(i) At the renegotiation stage, the initial security is renegotiated to an efficient piecewise-linear ambiguity-sharing security, given θ_E and θ_{I1}. That is, r_n^* satisfies equation ((19)) in Proposition 1, given θ_E, θ_{I1} and a^*, with λ increasing in D^*.

(ii) The entrepreneur's effort a^* is strictly lower than in the first-best problem given θ_E and θ_{I1}.

As shown in Example 1, the security that solves the first-best problem may be any one of convertible debt, levered equity, or unlevered equity, depending on the exact parameter values. Thus, the solution to the security design problem with renegotiation sees risky debt issued at date 0, with the debt being transformed by the addition of an equity component (in the case of levered equity or convertible debt) or being directly converted to equity (in the case of unlevered equity) at the renegotiation stage. Although in our model renegotiation is assumed to always occur, it is natural to think of renegotiation as being dependent on the outcomes that obtain between dates 0 and 0.5. In this case, the initial security itself may be interpreted as convertible debt — it starts out as debt, and either converts wholly or partially into equity at date 0.5, or at the minimum specifies a precise conversion option.

The solution here has some resemblance to the outcomes that result in stage financing when a firm raises money from venture capitalists in multiple stages. An important difference, of course, is that we do not model the need for additional investment at the new stage, so our analysis can be interpreted in terms of the security held by an initial investor who does not inject new capital at the next stage.

Our results here are similar to those of Dewatripont et al. (2003), with two differences. First, we assume that the value investors' derive from a security changes between dates 0 and 0.5, with new information leading to an increase in θ_I. Second, both the entrepreneur and the investors are ambiguity-averse in our model, creating a need for ambiguity-sharing. Of course, we exploit the fact that the Hansen and Sargent (2001) approach to model uncertainty results in functional forms that resemble CARA utility, so at that stage the outcome looks similar to what would obtain under risk-sharing. In contrast, Dewatripont et al. (2003) model the investors as risk-neutral and the entrepreneur as risk-averse, so that after renegotiation the entrepreneur obtains a constant wage. In practice, entrepreneurs tend to retain substantial control rights in the firm at each stage. This is captured in our model with the idea that the entrepreneur's final stake is also piecewise-linear in the cash flow.

6.2 Example 3

6.2.1 Simplified Model: Binary Outcomes

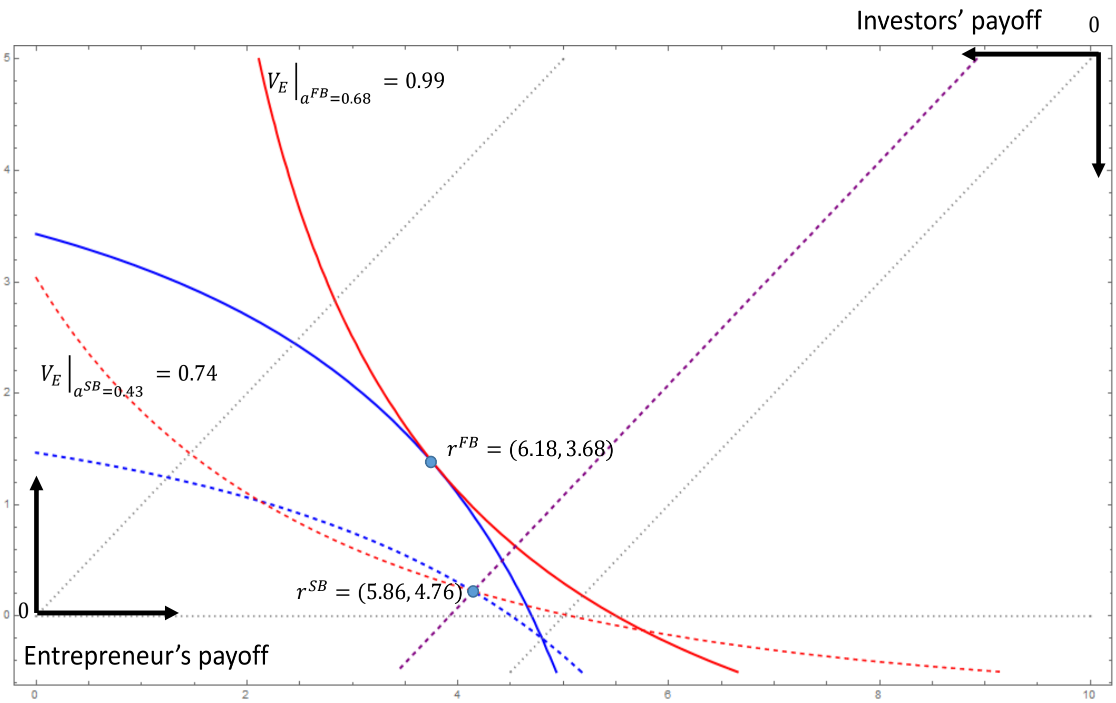
To explain the intuition behind the model, we first solve a simplified model with binary cash flows. The cash flow is either $x_g = 10$ or $x_b = 5$, with $\text{Prob}(x_g \mid a) = a$, where the action a lies in $[0, 1]$. Set $\psi(a) = 4\,a^2$ and $I = 5.2$. Let $\theta_E = \theta_{I0} = 4$, so that the entrepreneur and the investors share the same ambiguity aversion parameter at date 0. At the renegotiation stage, $\theta_{I1} = 10$.

As benchmarks, we first present the first- and second-best solutions when the contracting parties commit to not renegotiating, and letting the original contract stand. We assume that in this case, θ_I remains at 4 throughout. The Edgeworth box in Figure 6 illustrates the results. The 45° line on the left depicts constant wage contracts that perfectly insure the entrepreneur. Conversely, the rightmost 45° degree line segment above the horizontal line depicts risk-free debt contracts that perfectly insure the investor. The gray shaded parallelogram represents feasible contracts that satisfy the entrepreneur's and investors' limited liability and monotonicity conditions. Risky debt is represented the contracts on the horizontal axis to the left of the risk-free debt line. With risky debt, all cash flow is given to the investor in the low cash flow state, and the entrepreneur receives zero in this state. The entrepreneur prefers contracts to the northeast whereas the investor prefers those to the southwest.

When renegotiation is ruled out, the first-best security (the contract r^{FB} in the figure) provides optimal ambiguity sharing between the entrepreneur and the investor with ambiguity aversion parameter $\theta_{I0} = 4$. The second-best security is represented by r^{SB}; as expected, the need to provide incentives to the entrepreneur distorts ambiguity sharing.

Now, suppose regulation is allowed for. At the initial stage, the entrepreneur offers risky debt with face value $D^* = 5.34$, so that the investors payoff is D^* in the high cash flow state, and $x_b = 5$ in the low state. Conversely, the entrepreneur obtains 4.66 in the high state and 0 in the low state. At the renegotiation stage, the investor and the entrepreneur agree on an efficient ambiguity-sharing contract denoted by r^n.

Table 2 summarizes the results in the binary output case. Note that the entrepreneur's expected payoff is $V_E = -\theta_E \ln\left(\int_X e^{-(x-r(x))/\theta_E} f(x \mid a)dx\right) - \psi(a)$, as in equation (6). The entrepreneur's expected payoff with renegotiation is of course higher than the payoff in the second-best case without renegotiation. In our example, because θ_{I1} is substantially higher than θ_{I0}, the gains to trade from renegotiation are high, so that the entrepreneur's expected payoff with renegotiation in fact exceeds the payoff in the first-best case without renegotiation (in which θ_I remains unchanged



The solid blue line represents the investor's individual rationality constraint when the entrepreneur exerts first-best effort. The dashed blue line represents the same IR constraint under second-best effort. The solid red line and dashed red line represent the entrepreneurs' iso-utility curve under first- and second-best effort, respectively.

Figure 6: First-best and second-best securities without renegotiation

	Without renegotiation		With renegotiation
	First-best	Second-best	
Optimal contract	(6.18, 3.68)	(5.86, 4.76)	(6.79, 3.22)
Initial contract	-	-	(5.34,5)
Optimal effort	0.68	0.43	0.6
Entrepreneur's expected payoff	0.99	0.74	1.14

Table 2: Summary

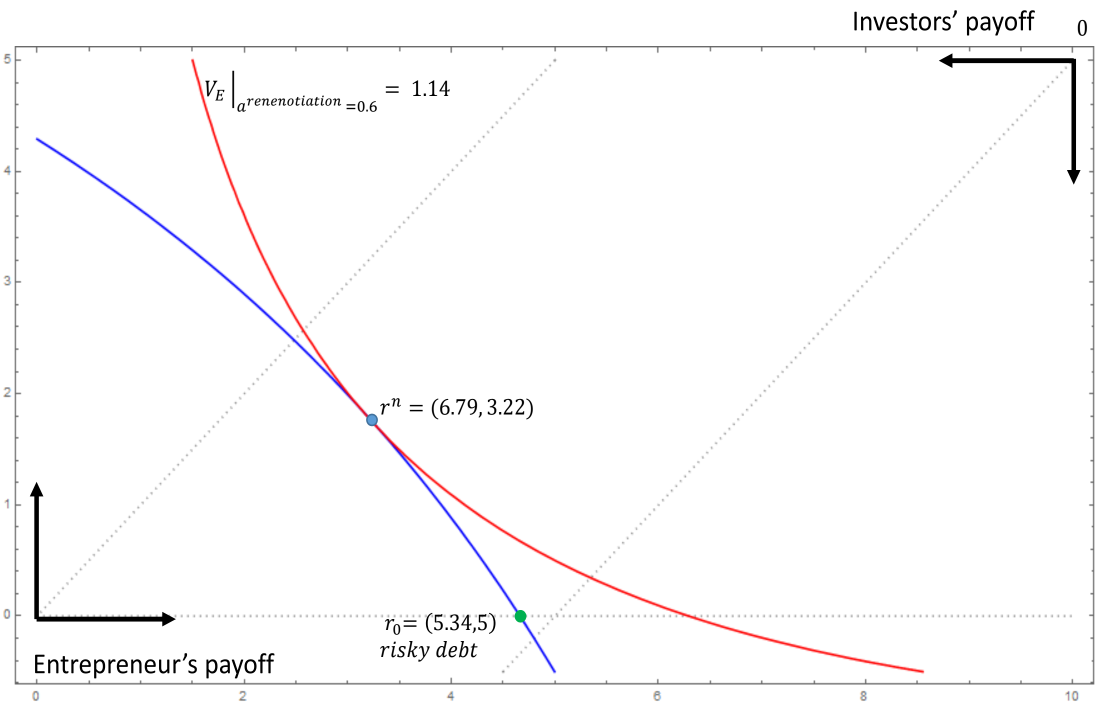
at 4).

6.2.2 The full model

Now, as in Examples 1 and 2, let the cash-flow set be $X = [0,1]$, and let the action set be $A = [0,1]$. Set $f(x \mid a) = (1-a)\cdot 1 + a(2x) = 1 + a(2x-1)$ and let $\psi(a) = \frac{1}{2}a^2$. Finally, let $I = 0.3$.

At the initial stage, we assume that the entrepreneur's robustness parameter is $\theta_E = 1$ and the investor's parameter is $\theta_I = 1.422$. At the renegotiation stage, θ_I increases from 1.422 to 4.

First, suppose there is no renegotiation, and the investors' ambiguity aversion parameter remains



At the initial stage, the entrepreneur offers risky debt with face value $D^* = 5.34$, so that the investors payoff is D^* in the high cash flow state, and $x_b = 5$ in the low state - that is, the entrepreneur obtains 4.66 in the high state and 0 in the low state. At the renegotiation stage, the investor and the entrepreneur agree on an efficient ambiguity-sharing contract denoted by r^n.

Figure 7: Optimal contract with renegotiation

$\theta_{I0} = 1.422$. The blue solid line in Figure 8 depicts the first-best security and the blue dashed line the second-best security in this case.

Next, suppose that θ_I increases to 4 at date 0.5, and suppose that renegotiation can occur. The optimal initial contract is now risky debt with face value $D^* = 0.361$. At the renegotiation, efficient ambiguity-sharing leads to the red solid line, which represents a levered equity contract.

	Without renegotiation		With renegotiation
	First-best	Second-best	
Optimal contract	Equity	Convertible debt	Levered equity
Initial contract	-	-	Risky debt
Optimal effort	0.168	0.082	0.12
Entrepreneur's expected payoff	0.1968	0.1926	0.2013

Table 3: Summary

Table 3 summarizes the optimal actions and the payoffs of the entrepreneur in the various cases we consider. Here too, we find that the large increase in θ_I generates sufficient gains from trade at

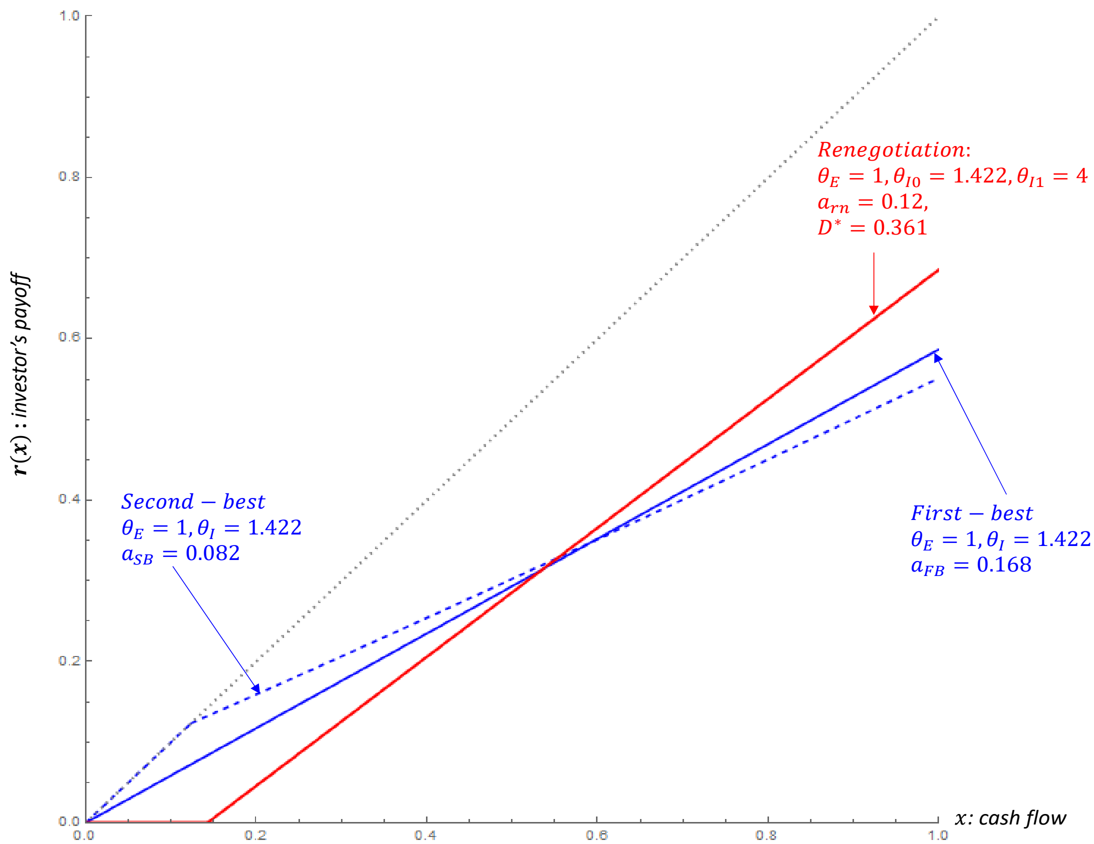


Figure 8: Optimal securities with ex-post renegotiation

the renegotiation stage to result in a higher expected payoff for the entrepreneur, as compared to the case in which θ_I is constant and no renegotiation can occur.

Recall that due to the entrepreneur's incentive compatibility condition, the second-best contract is non-linear (Proposition 2). Hence, it cannot be directly implemented with equity, even though it can have equity-like component. However, as we have shown, with renegotiation the optimal contracts are piecewise-linear at both the initial and renegotiation stages (debt and levered equity, respectively).

7 Conclusion

We extend the Innes (1990) model of a risk-neutral entrepreneur and investors to allow for ambiguity-aversion, using the multiplier preferences introduced by Hansen and Sargent (2001). Ambiguity aversion of both parties creates a benefit from ambiguity-sharing. If effort is directly contractible, the optimal contract features a security that directly includes an equity component, and is interpretable as either convertible debt or levered equity.

When the entrepreneur's action is not contractible, the optimal contract must be designed to provide incentives for effort. The investor now receives more cash in low cash flow states, compared to the security in the first-best contract. If the entrepreneur is sufficiently risk-averse, the contract resembles equity in high cash flow states, but has payments to the investors that are non-linear in the project cash flow. If the entrepreneur has a low degree of ambiguity aversion, the security can have non-monotone payments that decrease over some range of project cash flow. In this case, imposing monotonicity of the claims held by both entrepreneur and investors leads to the optimal security being either plain vanilla debt, or debt that with a conversion option in high cash flow states.

Over time, more information about the firm becomes available, and to the extent that multiplier preferences can also be represented as constraint preferences, it is reasonable to think of both parties become less ambiguity-averse in the future. We consider the implications of this for security design. For simplicity, we assume only the investors' ambiguity-aversion parameter changes, and allow for renegotiation of the initial contract following such a change. We show that renegotiation in this setting eliminates non-linear components in the second-best contract, because incentive compatibility is no longer issue at the renegotiation. As a result, convertible debt emerges as an optimal contract: the entrepreneur initially offers a risky debt contract, which is renegotiated to either explicitly convertible debt or levered equity.

Our results imply that ambiguity-sharing may underlie the design of venture capital contracts, which generally feature either an equity component or convertibility. While there are many theories that lead to the optimality of equity or convertible debt, our model provides a parsimonious and plausible explanation for the existence of such contracts.

A Appendix: Proofs

Proof of Lemma 1

First, consider the IR constraint in equation (11). Dividing throughout by $-\theta_I$, we have

$$\ln\left(\int_X e^{-\frac{r(x)}{\theta_I}} f(x \mid a)dx\right) \leq -\frac{I}{\theta_I} = \ln\left(e^{-\frac{I}{\theta_I}}\right).$$

Taking the exponential of both sides yields the constraint (IR2) exhibited in equation (15).

Next, consider the IC constraint in equation (12). Multiply throughout by $-\frac{\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x|a)dx}{\theta_E}$ to obtain

$$\int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x \mid a)dx + \frac{\psi'(a)}{\theta_E}\int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x \mid a)dx = 0,$$

which is the constraint (IC2) exhibited in equation (16).

Now, observe that the limited liability constraints are identical in problems P1 and P2. As the IR and IC constraints are also equivalent across these problems, the feasible sets of (r, a) are identical in both problems.

Finally, consider the objective function in problem P1, as exhibited in equation (10). Denote $\Phi(r, a) = -\theta_E \ln\left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx\right) - \psi(a)$. Then,

$$\begin{aligned}
\Phi(r, a) &= -\theta_E\left[\ln\left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx\right) + \frac{\psi(a)}{\theta_E}\right] \\
&= -\theta_E\left[\ln\left(\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx\right) + \ln\left(e^{\frac{\psi(a)}{\theta_E}}\right)\right] \\
&= -\theta_E \ln\left(e^{\frac{\psi(a)}{\theta_E}}\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx\right). \qquad (35)
\end{aligned}$$

Now, maximizing $\Phi(r, a)$ is equivalent to minimizing $-\frac{1}{\theta_E}\Phi(r, a) = \ln\left(e^{\frac{\psi(a)}{\theta_E}}\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx\right)$. Finally, minimizing the last expression is equivalent to minimizing its exponential, $e^{\frac{\psi(a)}{\theta_E}}\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a)dx$. The latter is the object being minimized in equation (14) in Problem P2.

The problems P1 and P2 are therefore equivalent, and must have the same solution sets. ■

Proof of Proposition 1

(i) Note that, as $I > 0$, the investors' IR constraint can only be satisfied if $r(x) > 0$ over some set of positive measure, $Y \subseteq X$. Now, suppose the IR constraint is slack and the optimal security is $\tilde{r}(x)$, so that $\int_X e^{-\frac{\tilde{r}(x)}{\theta_I}} f(x \mid a)dx < e^{-\frac{I}{\theta_I}}$. For some $\epsilon > 0$, set $\hat{r}(x) = \tilde{r}(x) - \epsilon$ if $x \in Y$ and $\hat{r}(x) = \tilde{r}(x)$ if $x \notin Y$. As $\tilde{r}(x) > 0$ for all $x \in Y$, there exists some $\epsilon > 0$ such that $\hat{r}(x) \geq 0$ for

$x \in Y$ and $\int_X e^{-\frac{\hat{r}(x)}{\theta_I}} f(x \mid a)dx < e^{-\frac{I}{\theta_I}}$. It is immediate that the security $\hat{r}(x)$ yields a higher payoff to the entrepreneur than $\tilde{r}(x)$, so that $\tilde{r}(x)$ cannot be an optimal security.

(ii) Optimize the Lagrangian pointwise with respect to $r(x)$. At a fixed value of x, the first-order condition $\frac{\partial \mathcal{L}_f}{\partial r} = 0$ yields

$$\begin{aligned} \frac{e^{\frac{\psi(a)}{\theta_E}}}{\theta_E} e^{-\frac{x-r(x)}{\theta_E}} f(x \mid a) - \frac{\lambda}{\theta_I} e^{-\frac{r(x)}{\theta_I}} f(x \mid a) - \underline{\gamma}_x + \overline{\gamma}_x &= 0 \\ \left[e^{-\frac{x-r(x)}{\theta_E} + \frac{\psi(a)}{\theta_E}} - \frac{\lambda \theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I}} \right] \frac{f(x \mid a)}{\theta_E} &= \underline{\gamma}_x - \overline{\gamma}_x. \end{aligned} \tag{36}$$

Now, there are three cases to consider.

Case 1: $\underline{\gamma}_x > 0$. Then, $r(x) = 0$ by complementary slackness, so it follows that $\overline{\gamma}_x = 0$. Equation (36) reduces to

$$\left[e^{\frac{-x+\psi(a)}{\theta_E}} - \frac{\lambda \theta_E}{\theta_I} \right] \frac{f(x \mid a)}{\theta_E} = \underline{\gamma}_x. \tag{37}$$

As $\frac{f(x|a)}{\theta_E} > 0$, it follows that $e^{\frac{-x+\psi(a)}{\theta_E}} > \frac{\lambda\theta_E}{\theta_I}$. Taking natural logs on both sides and rearranging, we have

$$x < \psi(a) - \theta_E \ln \left(\frac{\lambda \theta_E}{\theta_I} \right). \tag{38}$$

As $x > 0$, there exist values of x for which this case is feasible only if $\ln\left(\frac{\lambda\theta_E}{\theta_I}\right) < \frac{\psi(a)}{\theta_E}$.

Case 2: $\overline{\gamma}_x > 0$. Then, $r(x) = x$ by complementary slackness, so it follows that $\underline{\gamma}_x = 0$. Equation (36) reduces to

$$\left[e^{\frac{\psi(a)}{\theta_E}} - \frac{\lambda \theta_E}{\theta_I} e^{-\frac{x}{\theta_I}} \right] \frac{f(x \mid a)}{\theta_E} = -\overline{\gamma}_x. \tag{39}$$

As $\frac{f(x|a)}{\theta_E} > 0$, it follows that $e^{\frac{\psi(a)}{\theta_E}} < \frac{\lambda\theta_E}{\theta_I} e^{-\frac{x}{\theta_I}}$. Taking natural logs on both sides and rearranging, we have

$$x < \theta_I \left[\ln \left(\frac{\lambda \theta_E}{\theta_I} \right) - \frac{\psi(a)}{\theta_E} \right]. \tag{40}$$

As $x > 0$, there exist values of x for which this case is feasible only if $\ln\left(\frac{\lambda\theta_E}{\theta_I}\right) < \frac{\psi(a)}{\theta_E}$.

Case 3: $\overline{\gamma}_x = \underline{\gamma}_x = 0$. Then, $r(x) \in (0, x)$. Here, equation (36) reduces to

$$\left[e^{-\frac{x-r(x)}{\theta_E}+\frac{\psi(a)}{\theta_E}} - \frac{\lambda\theta_E}{\theta_I}e^{-\frac{r(x)}{\theta_I}}\right]\frac{f(x \mid a)}{\theta_E} \quad = \quad 0. \tag{41}$$

As $\frac{f(x|a)}{\theta_E} > 0$, it must be that $e^{-\frac{x-r(x)}{\theta_E}+\frac{\psi(a)}{\theta_E}} = \frac{\lambda\theta_E}{\theta_I}e^{-\frac{r(x)}{\theta_I}}$. Taking natural logs on both sides, we have

$$\begin{aligned} -\frac{x}{\theta_E} + \frac{r(x)}{\theta_E} + \frac{\psi(a)}{\theta_E} \quad &= \quad \ln\left(\frac{\lambda\theta_E}{\theta_I}\right) - \frac{r(x)}{\theta_I} \\ r(x) \quad = \quad \frac{\theta_I}{\theta_I+\theta_E}x + \frac{\theta_I\theta_E}{\theta_I+\theta_E}&\left(\ln\left(\frac{\lambda\theta_E}{\theta_I}\right) - \frac{\psi(a)}{\theta_E}\right). \end{aligned} \tag{42}$$

Now, let λ_f and a_f be the values of λ and a when the Lagrangian has been optimized. Denote the RHS of (42), evaluated at $\lambda = \lambda_f$ and $a = a_f$, by $\hat{r}(x)$. It follows that if $\hat{r}(x) \in (0, x)$, then $r_f(x) = \hat{r}(x)$. If $\hat{r}(x) < 0$, then $r_f(x) = 0$, so that we are in Case 1. Note that this case can occur only if $\ln\left(\frac{\lambda\theta_E}{\theta_I}\right) < \frac{\psi(a)}{\theta_E}$. Finally, if $\hat{r}_x > x$, then $r_f(x) = x$, putting us in Case 2. Note that this case can occur only if $\ln\left(\frac{\lambda\theta_E}{\theta_I}\right) > \frac{\psi(a)}{\theta_E}$.

Substitute $\frac{\psi(a)}{\theta_E} = \ln e^{\frac{\psi(a)}{\theta_E}}$. Then, the contract in the statement of the proposition, in equation (19), succinctly describes the three cases. ■

Proof of Proposition 2

We prove part (ii) first and then part (i).

(ii) The proof of part (ii) closely mirrors the proof of Propostion 1 (ii).

Optimize the Lagrangian pointwise with respect to $r(x)$. At a fixed value of x, the first-order condition $\frac{\partial\mathcal{L}}{\partial r} = 0$ yields

$$\begin{aligned} &\frac{e^{\frac{\psi(a)}{\theta_E}}}{\theta_E}e^{-\frac{x-r(x)}{\theta_E}}f(x \mid a) - \frac{\lambda}{\theta_I}e^{-\frac{r(x)}{\theta_I}}f(x \mid a) \\ &\qquad +\frac{\mu}{\theta_E}\left(e^{-\frac{x-r(x)}{\theta_E}}f_a(x \mid a) + \frac{\psi'(a)}{\theta_E}e^{-\frac{x-r(x)}{\theta_E}}f(x \mid a)\right) - \underline{\gamma}_x + \overline{\gamma}_x \quad = \quad 0 \\ &e^{-\frac{x-r(x)}{\theta_E}}\left[e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right) - \frac{\lambda\theta_E}{\theta_I}e^{-\frac{r(x)}{\theta_I}+\frac{x-r(x)}{\theta_E}}\right]\frac{f(x \mid a)}{\theta_E} \quad = \quad \underline{\gamma}_x - \overline{\gamma}_x. \end{aligned} \tag{43}$$

Now, there are three cases to consider.

Case 1: $\underline{\gamma}_x > 0$. Then, $r(x) = 0$ by complementary slackness, so it follows that $\overline{\gamma}_x = 0$. Equation (43) reduces to

$$e^{-\frac{x}{\theta_E}}\left[e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right) - \frac{\lambda\theta_E}{\theta_I}e^{\frac{x}{\theta_E}}\right]\frac{f(x \mid a)}{\theta_E} \quad = \quad \underline{\gamma}_x. \tag{44}$$

As $e^{-\frac{x}{\theta_E}}$, $f(x \mid a)$, and θ_E are all strictly positive, it follows that

$$e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right) \quad > \quad \frac{\lambda\theta_E}{\theta_I} e^{\frac{x}{\theta_E}}. \tag{45}$$

Taking natural logs on both sides, we have

$$\ln\left(e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right)\right) \quad > \quad \ln\frac{\lambda\theta_E}{\theta_I} + \frac{x}{\theta_E}. \tag{46}$$

Recall that, by MLRP, $\frac{f_a(x|a)}{f(x|a)}$ is strictly increasing in x. Therefore, both sides of the last equation are strictly increasing in x. Therefore, the equation $\ln\left(e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x|a)}{f(x|a)} + \frac{\psi'(a)}{\theta_E}\right)\right) = \ln\frac{\lambda\theta_E}{\theta_I} + \frac{x}{\theta_E}$ can have zero or multiple roots, depending on parameters.

Case 2: $\overline{\gamma}_x > 0$. Then, $r(x) = x$ by complementary slackness, so it follows that $\underline{\gamma}_x = 0$. Equation (43) reduces to

$$\left[e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right) - \frac{\lambda\theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I}}\right] \frac{f(x \mid a)}{\theta_E} \quad = \quad -\overline{\gamma}_x. \tag{47}$$

As $\frac{f(x|a)}{\theta_E} > 0$, it follows that

$$\begin{aligned} e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right) \quad &< \quad \frac{\lambda\theta_E}{\theta_I} e^{-\frac{x}{\theta_I}} \\ \ln\left(e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right)\right) \quad &< \quad \ln\frac{\lambda\theta_E}{\theta_I} - \frac{x}{\theta_I}. \end{aligned} \tag{48}$$

The LHS of the last equation is strictly increasing in x, and the RHS is strictly decreasing. Therefore, either (a) the inequality is violated for all $x \geq 0$, or (b) there exists a threshold $\hat{x}$ such that the inequality holds for $x \leq \hat{x}$.

Case 3: $\underline{\gamma}_x = \overline{\gamma}_x = 0$. Then, $r(x) \in (0, x)$. Here, equation (43) reduces to

$$e^{-\frac{x-r(x)}{\theta_E}}\left[e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right) - \frac{\lambda\theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I} + \frac{x-r(x)}{\theta_E}}\right] \frac{f(x \mid a)}{\theta_E} \quad = \quad 0, \tag{49}$$

which implies that

$$e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right) \quad = \quad \frac{\lambda\theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I} + \frac{x-r(x)}{\theta_E}} \tag{50}$$

$$\ln\left(e^{\frac{\psi(a)}{\theta_E}} + \mu\left(\frac{f_a(x \mid a)}{f(x \mid a)} + \frac{\psi'(a)}{\theta_E}\right)\right) \quad = \quad \ln\frac{\lambda\theta_E}{\theta_I} - \frac{r(x)}{\theta_I} + \frac{x - r(x)}{\theta_E}. \tag{51}$$

The last equation directly implies that

$$r(x)=\frac{\theta_I}{\theta_I+\theta_E}x+\frac{\theta_I\theta_E}{\theta_I+\theta_E}\left\{\ln\frac{\lambda\theta_E}{\theta_I}-\ln\left(e^{\frac{\psi(a)}{\theta_E}}+\mu\left(\frac{f_a(x\mid a)}{f(x\mid a)}+\frac{\psi'(a)}{\theta_E}\right)\right)\right\}. \tag{52}$$

Now, let a^*,λ^*, and μ^* denote the values of the respective variables when the Lagrangian has been optimized. Let $\hat{r}(x)$ denote the RHS of equation (52). It follows that $r^*(x)=\hat{r}(x)$ when $\hat{r}(x)\in[0,x]$, $r^*(x)=0$ when $\hat{r}(x)<0$, and $r^*(x)=x$ when $\hat{r}(x)>x$. The statement of part (ii) describes these possibilities in a more succinct manner.

(i) Suppose that the IR constraint does not bind, so that $\lambda^*=0$. Consider the expression for $\hat{r}(x)$ in equation (52). As $\lambda\to 0$, regardless of the value of x, the term $\ln\frac{\lambda\theta_E}{\theta_I}\to-\infty$, so it follows that $\hat{r}(x)<0$ and $r(x)=0$. However, if $r(x)=0$ for all x, the IR constraint is trivially violated, so we have a contradiction. Therefore, the IR constraint must bind. ■

Proof of Proposition 3

As in the proof of Proposition 2, we first show part (ii).

Denote $\rho=r'(x)$. For any given a, the corresponding Hamiltonian (or point-wise Lagrangian) is

$$\begin{aligned}H(x,\lambda,\mu,r(\cdot),\rho(\cdot))&=e^{\frac{\psi(a)}{\theta_E}}e^{\frac{-(x-r(x))}{\theta_E}}f(x|a)+\lambda\left(e^{-\frac{r(x)}{\theta_I}}f(x|a)-e^{-I/\theta_I}\right)\\&\quad+\mu\left(\frac{\psi'(a)}{\theta_E}e^{-\frac{(x-r(x))}{\theta_E}}f(x|a)+e^{-\frac{(x-r(x))}{\theta_E}}f_a(x|a)\right)+\xi(x)\rho(x),\end{aligned} \tag{53}$$

where we temporarily suppress the limited liability constraints $0\le r(x)\le x$ and the constraint $\rho(x)\le 1$. Further, ξ is the costate variable associated with $\rho=r'$.

Let $\bar{r}(x)$ be the optimal security given that condition (M) has been imposed. By Pontryagin's minimum principle, the necessary conditions for an optimum $(\bar{r}(x),\rho(x))$ are:

(i) $\rho(x)=\text{argmin}_{0\le\tilde{\rho}(x)}H(x,\lambda,\mu,\bar{r}(\cdot),\tilde{\rho}(\cdot))$.

(ii) The costate variable associated with $\rho(x)$ satisfies

$$\xi'(x)\;=\;-\frac{\partial H}{\partial r}(x). \tag{54}$$

(iii) Since $r(0)=0$ (by limited liability for both investors and entrepreneur), but $r(\overline{x})$ can lie in the range $[0,\overline{x}]$, the transversality condition of the costate variable is $0=\xi(\overline{x})=-\int_0^{\overline{x}}\frac{\partial H}{\partial r}(x)dx$.

The optimality condition with respect to the control $\rho(x)$ is that, for all x,

$$\rho(x)=\underset{0\le\rho(x)\le 1}{\text{argmin}}\;H(x,\lambda,\mu,r(\cdot),\rho(\cdot))=\underset{0\le\rho(x)\le 1}{\text{argmin}}\;\xi(x)\rho(x) \tag{55}$$

That is,

$$\rho(x) = \begin{cases} 0 & \text{if } \xi(x) > 0 \\ 1 & \text{if } \xi(x) < 0 \\ (0,1) & \text{if } \xi(x) = 0 \end{cases} \tag{56}$$

The following cases emerge.

Case 1: $\xi(x) = 0$ over a range of positive measure, so $\xi'(x) = 0$ over this range. From equation (54), we have

$$\frac{\partial H}{\partial r}(x) = e^{\frac{-w(x)}{\theta_E}} f(x|a)\left(\frac{1}{\theta_E}e^{\frac{\psi(a)}{\theta_E}} - \frac{\lambda}{\theta_I}e^{-\frac{r(x)}{\theta_I}+\frac{x-r(x)}{\theta_E}} + \mu\left(\frac{\psi'(a)}{\theta_E^2} + \frac{1}{\theta_E}\frac{f_a(x|a)}{f(x|a)}\right)\right) = 0 \tag{57}$$

Notice that this last equation coincides with the optimality condition of the security $r(x)$ in the second-best contract when neither limited liability condition binds. Observe that in this case $r(x)$ is strictly increasing: $r'(x) = \rho(x) \in (0,1)$.

Case 2: $\xi(x) > 0$ over some range of positive measure. Then $\rho(x) = r'(x) = 0$ over this range, which implies $r(x) = c_0$, for some constant $c_0 \in \mathbb{R}$. From equation (54), we have

$$e^{\frac{-(x-c_0)}{\theta_E}} f(x|a)\left(\frac{\lambda}{\theta_I}e^{\frac{x}{\theta_E}-\frac{\theta_E+\theta_I}{\theta_I\theta_E}c_0} - \mu\frac{1}{\theta_E}\frac{f_a(x|a)}{f(x|a)} - \frac{1}{\theta_E}e^{\frac{\psi(a)}{\theta_E}} - \mu\frac{\psi'(a)}{\theta_E^2}\right) = \xi'(x) \tag{58}$$

Case 3: $\xi(x) < 0$ over some range of positive measure. Then $\rho(x) = r'(x) = 1$, which implies $r(x) = x - \hat{c}_0$ and $w(x) = \hat{c}_0$ for some $\hat{c}_0 \in \mathbb{R}$. From (54), we have

$$e^{-\frac{\hat{c}_0}{\theta_E}} f(x|a)\left(\frac{\lambda}{\theta_I}e^{-\left(\frac{x}{\theta_I}-\frac{\theta_I+\theta_E}{\theta_E\theta_I}\hat{c}_0\right)} - \mu\frac{1}{\theta_E}\frac{f_a(x|a)}{f(x|a)} - \frac{1}{\theta_E}e^{\frac{\psi(a)}{\theta_E}} - \mu\frac{\psi'(a)}{\theta_E^2}\right) = \xi'(x) \tag{59}$$

Observe that the mapping

$$x \mapsto \frac{\lambda}{\theta_I}e^{-\left(\frac{x}{\theta_I}-\frac{\theta_I+\theta_E}{\theta_E\theta_I}\hat{c}_0\right)} - \mu\frac{1}{\theta_E}\frac{f_a(x|a)}{f(x|a)} - \frac{1}{\theta_E}e^{\frac{\psi(a)}{\theta_E}} - \mu\frac{\psi'(a)}{\theta_E^2} \tag{60}$$

is strictly decreasing in x.

Given the parameters (θ_E, θ_I), the optimal values of (a, λ, μ) must satisfy one of the following cases.

(i) Suppose that $\xi(0) > 0$. Then, as $r(0) = 0$ by limited liability for both investors and entrepreneur, it follows that $c_0 = 0$. Then, the equation $\xi(x) = 0$ must have a solution in $[0, \overline{x}]$. Otherwise, $\xi(x) > 0$ for all $x \in [0, \overline{x}]$, which implies $r(x) = 0$ for all $x \in [0, \overline{x}]$, which violates the investors' IR constraint. Let $\hat{x}_0(a, \lambda, \mu)$ be such solution. Then, it must be that $\xi(x) = 0$

for all $x \in [\hat{x}_0, \overline{x}]$. To see this, suppose that $\xi(x) < 0$ for some $\tilde{x} \in [\hat{x}_0, \overline{x}]$. As the mapping in equation (60) is strictly decreasing, it must be that $\xi'(x) < 0$ for all $x > \tilde{x}$. However, in this case, we have $\xi(\overline{x}) < 0$, which violates the boundary condition $\xi(\overline{x}) = 0$. Therefore, $\xi(x) = 0$ on $[\hat{x}_0, \overline{x}]$ and we have the situation in equation (57).

(ii) Suppose that $\xi(0) < 0$. Then, as $r(0) = 0$, it follows that $\hat{c}_0 = 0$. Also, it must be that $\xi'(0) > 0$. Otherwise, we have $\xi'(0) \leq 0$ for all $x \in [0, \overline{x}]$, which would lead to a violation of the boundary condition $\xi(\overline{x}) = 0$. With $\xi(0) < 0$ and $\xi'(0) > 0$, the function $\xi(x)$ is an increasing and concave function in x on $[0, \hat{x}_0]$ where $\hat{x}_0 = \hat{x}_0(a, \lambda, \mu)$ is a solution of the equation $\xi(x) = 0$. Further, it follows that $\xi(x) \geq 0$ for all $x \in [\hat{x}_0, \overline{x}]$. Then,

   (a) Suppose that $\xi'(x) < 0$ for all $x \in [\hat{x}_0, \overline{x}]$. Then $\xi(x) = 0$ for all $x \in [\hat{x}_0, \overline{x}]$. Thus $r(x) = r^*(x)$ on $x \in [\hat{x}_0, \overline{x}]$.

   (b) Suppose that $\xi'(x) > 0$ for some $x \in [\hat{x}_0, \hat{x}_1]$ and $\xi'(x) < 0$ for some $x \in [\hat{x}_1, \hat{x}_2]$, where $\xi'(\hat{x}_1) = 0$. Then, $\xi(x) > 0$ for $x \in [\hat{x}_1, \hat{x}_2]$. Hence $r(x)$ is a constant for $x \in [\hat{x}_1, \hat{x}_2]$. The decreasing mapping in equation (60) and the transversality condition $\xi(\overline{x}) = 0$ ensure that $\xi(x) = 0$ for $x \in [\hat{x}_2, \overline{x}]$. That is, $r(x) = r^*(x)$ on $[\hat{x}_2, \overline{x}]$. Finally, if $\hat{x}_2 = \overline{x}$, then the security is standard debt.

The function $r^M(x)$ in equation (25) encapsulates these various cases.

The proof of part (i) now completely mirrors the proof of Proposition 2, part (i). ■

Relation between multiplier preferences and constraint preferences

Define the constraint preference problem

$$J(\eta) \triangleq \min_{p \in \Delta(q)} \left\{ \int u(x)p(x)dx : \, R(p||q) \leq \eta \right\} \tag{61}$$

Next, construct the Lagrangian associated with $J(\eta)$:

$$\mathcal{L} = \int u(x)p(x)dx + \theta \left(R(p||q) - \eta \right) \tag{62}$$

Now, define the Lagrangian dual function to $J(\eta)$ as

$$L(\theta) \triangleq \min_{p \in \Delta(q)} \left\{ \int \left(u(x) - \theta \, \eta \right) p(x)dx + \theta R(p||q) \right\} \tag{63}$$

Assume that strong duality holds. Then, we have

Lemma 3. *Suppose the minimum of $J(\eta)$ is finite. Then $\exists \theta > 0$ such that:*

$$J(\eta) = \max_{\theta > 0} L(\theta) \tag{64}$$

Proof. See Chapter 6 in Luenberger (1969). □

Now, by the representation Lemma in Dupuis and Ellis (1997), we have

$$L(\theta) \;=\; \min_{p\in\Delta(q)}\left\{\int(u(x)-\theta\,\eta)\,p(x)dx+\theta R(p||q)\right\}=-\theta\log\left(\int\exp\left(-\frac{u(x)}{\theta}\right)q(x)dx\right)-\theta\cdot\eta$$

The optimal θ satisfies the first-order condition:

$$-\log\left(\int e^{-u(x)/\theta}q(x)dx\right)+\frac{1}{\theta}\frac{\int\exp\left(-u(x)/\theta\right)u(x)q(x)dx}{\int\exp\left(-u(x)/\theta\right)q(x)dx}-\eta=0$$

Therefore, the optimal $\theta^*(\eta)$ depends on η: a change in η leads to θ. □

Proof of Lemma 2

Proof. Fix a, so that $T(\cdot)$ varies only with r_0. By Proposition 1 and 2 in Luenberger (1969) Section 8.3, it is sufficient observe that the objective functional and the constraint functional in (33) are convex in $r_n(\cdot)\in C[0,\overline{x}]$. Then, $T(\cdot)$ is convex and decreasing in $\hat{u}$. □

Proof of Proposition 4

Proof. (i) Suppose that $r_0(\cdot)$ is an arbitrary initial contract that satisfies monotonicity; that is, $r'(x)\in[0,1]$ for all x. Then,

$$\begin{aligned}\frac{\partial}{\partial a}\hat{u}(\cdot) &= \int_X e^{-\frac{r_0(x)}{\theta_{I1}}}f_a(x|a)dx=\int_X e^{-\frac{r_0(x)}{\theta_{I1}}}\frac{f_a(x|a)}{f(x|a)}f(x|a)dx\\ &= \mathbb{E}_{x\sim f(\cdot|a)}\left[e^{-\frac{r_0(x)}{\theta_{I1}}}\cdot\left(\frac{f_a(x|a)}{f(x|a)}\right)\right]\overset{(a)}{=}\mathbb{COV}\left[e^{-\frac{r_0(x)}{\theta_{I1}}},\left(\frac{f_a(x|a)}{f(x|a)}\right)\right]\overset{(b)}{\leq}0.\end{aligned}\tag{65}$$

Here, the equality (a) is valid because

$$\mathbb{E}\left[\frac{f_a(x|a)}{f(x|a)}\right]=\int_X\frac{f_a(x)}{f(x)}f(x)dx=\int_X f_a(x)dx=\frac{\partial}{\partial a}\left(\int_X f(x)dx\right)=0,\tag{66}$$

where the derivative and the integral may be interchanged because the domain X is compact. Further, the inequality (b) is valid because the log-likelihood ratio $\left(\frac{f_a(x|a)}{f(x|a)}\right)$ is increasing in x by MLRP, and the function $x\mapsto e^{-\frac{r_0(x)}{\theta_{I1}}}$ is decreasing by the monotonicity assumption on $r_0(x)$.

Now, by Lemma 2, keeping a fixed, $T(\cdot)$ is strictly decreasing and convex in $\hat{u}$. Thus, an optimal initial contract must be set to minimize $\left|\frac{\partial\hat{u}}{\partial a}\right|$. Consider a risky debt contract with face value D, denoted by $\delta(x)=\min\{x,D\}$. Let $r(\cdot)\in\mathcal{C}$ be any non-debt contract that provides the same expected utility as $\delta(\cdot)$, so that

$$\int e^{-\frac{r(x)}{\theta_{I1}}} f(x \mid a)dx = \int e^{-\frac{\delta(x)}{\theta_{I1}}} f(x \mid a)dx. \tag{67}$$

Now, define $\beta(x) := e^{-r(x)/\theta_{I1}} - e^{-\delta(x)/\theta_{I1}}$. It is immediate that β is decreasing in x. Hence, there exists a unique $\hat{x} \in (0, \overline{x})$ such that $\beta(x) \geq 0$ for $x \in [0, \hat{x}]$ and $\beta(x) \leq 0$ for $x \in (\hat{x}, \overline{x}]$. By the MLRP assumption and the fact that $\int \beta(x) f(x|a)dx = 0$, we have

$$\begin{aligned}
\left|\frac{\partial}{\partial a}\hat{u}(r,a)\right| - \left|\frac{\partial}{\partial a}\hat{u}(\delta,a)\right| &= -\int_0^{\overline{x}} \beta(x) \cdot \left(\frac{f_a(x|a)}{f(x|a)}\right) f(x|a)dx \\
&= \int_0^{\hat{x}} -\beta(x) \left(\frac{f_a(x|a)}{f(x|a)}\right) f(x|a)dx + \int_{\hat{x}}^{\overline{x}} -\beta(x) \left(\frac{f_a(x|a)}{f(x|a)}\right) f(x|a)dx \\
&> \int_0^{\hat{x}} -\beta(x) \left(\frac{f_a(\hat{x}|a)}{f(\hat{x}|a)}\right) f(x|a)dx + \int_{\hat{x}}^{\overline{x}} -\beta(x) \left(\frac{f_a(\hat{x}|a)}{f(\hat{x}|a)}\right) f(x|a)dx \\
&= \left(\frac{f_a(\hat{x}|a)}{f(\hat{x}|a)}\right) \int_0^{\overline{x}} -\beta(x) f(x|a)dx = 0
\end{aligned} \tag{68}$$

That is over all contracts that provide a given utility $\hat{u}$ with fixed a, risky debt minimizes $\left|\frac{\partial \hat{u}}{\partial a}\right|$. Hence, the optimal security at date 0 is a risky debt security for some face value D^*.

Now, consider the renegotiation stage. At this stage, effort is sunk at a^*. It is immediate that the entrepreneur's offer will be an efficient ambiguity-sharing contract, which satisfies (19) in Proposition 1 given a^*, θ_E and θ_{I1}, and where λ increases with D^*.

(ii) As $\frac{\partial \hat{u}}{\partial a} < 0$ for any feasible contract $r \in \mathcal{C}$ and $T(\cdot)$ is strictly decreasing in $\hat{u}$, it follows immediately that the optimal effort is strictly less than in the first-best problem given θ_E and θ_{I1} (in which the externality induced by effort on $\hat{u}$ is absent). $\square$

References

Andrey Malenko, Anton Tsoy (20018), *Asymmetric Information and Security Design under Knightian Uncertainty*, Working Paper.

Bewley, Truman F. (1989), Market innovation and entrepreneurship: A Knightian view, Cowles Foundation Discussion Paper 905.

Dewatripont, Mathias, Patrick Legros and Steven A. Matthews (2003), 'Moral hazard and capital structure dynamics', *Journal of the European Economic Association* (4), 890–930.

Dicks, David and Paolo Fulghieri (2015), 'Ambiguity, disagreement, and allocation of control in firms', *Working Paper, University of North Carolina* .

Dicks, David and Paolo Fulghieri (2016), 'Innovation waves, investor sentiment, and mergers', *Working Paper, University of North Carolina* .

Dupuis, Paul and Richard S. Ellis (1997), *A Weak Convergence Approach to the Theory of Large Deviations*, Wiley.

Ellsberg, Daniel (1961), 'Risk, ambiguity, and the Savage axioms', *The Quarterly Journal of Economics* **75**(4), 643–669.

Garlappi, Lorenzo, Ron Giammarino and Ali Lazrak (2015), 'ambiguity and the corporation: Group decisions, time inconsistency and underinvestment', Working paper, SSRN.

Gilboa, Itzhak and David Schmeidler (1989), 'Maxmin expected utility with non-unique prior', *Journal of Mathematical Economics* **18**(2), 141–153.

Gompers, Paul and Josh Lerner (2001), 'The venture capital revolution', *Journal of Economic Perspectives* **15**(2), 145–168.

Hansen, Lars P. and Thomas J. Sargent (2001), 'Robust control and model uncertainty', *American Economic Review* **91**(2), 60–66.

Hellmann, Thomas (2006), 'IPOs, acquisitions, and the use of convertible securities in venture capital', *Journal of Financial Economics* **81**(3), 649–679.

Hermalin, Benjamin E. and Michael L. Katz (1991), 'Moral hazard and verifiability: The effects of renegotiation in agency', *Econometrica* .

Innes, Robert D. (1990), 'Limited liability and incentive contracting with ex-ante action choices', *Journal of Economic Theory* **52**(1), 45–67.

Kaplan, Steven and Per Strömberg (2003), 'Financial contracting theory meets the real world: An empirical analysis of venture capital contracts', *Review of Economic Studies* **70**(2), 281–315.

Klibanoff, Peter, Massimo Marinacci and Sujoy Mukerji (2005), 'A smooth model of decision making under ambiguity', *Econometrica* **73**(6), 1849–1892.

Knight, Frank H. (1921), *Risk, Uncertainty and Profit*, Dover Classics.

Luenberger, David (1969), *Optimization by Vector Space Methods*, Wiley-Interscience.

Maccheroni, Fabio, Massimo Marinacci and Aldo Rustichini (2006), 'Ambiguity aversion, robustness, and the variational representation of preferences', *Econometrica* **74**(6), 1447–1498.

Miao, Jianjun and Alejandro Rivera (2016), 'Robust contracts in continuous time', *Econometrica* p. Forthcoming.

Ortner, Juan and Martin C. Schmalz (2016), Disagreement in optimal security design, SSRN working paper.

Ozerturk, Saltuk (2008), 'Risk sharing, risk shifting and the role of convertible debt', *Journal of Mathematical Economics* **44**(11), 1257–1265.

Schmidt, Klaus M. (2003), 'Convertible securities and venture capital finance', *Journal of Finance* **58**(3), 1139–1166.

Stein, Jeremy (1992), 'Convertible bonds as backdoor equity financing', *Journal of Financial Economics* **32**, 3–21.

Strzalecki, Tomasz (2011), 'Axiomatic foundations of multiplier preferences', *Econometrica* **79**(1), 47–73.

Szydlowski, Martin (2012), Ambiguity in dynamic contracts, Working paper, SSRN.

Wilson, Robert (1968), 'The theory of syndicates', *Econometrica* **36**, 119–132.