

# Robust Security Design\*

Seokwoo Lee<sup>†</sup>      Uday Rajan<sup>‡</sup>

December, 2019

## Abstract

We consider the optimal contract between an entrepreneur and investors in a moral hazard model when both parties have limited liability, are risk-neutral toward cash flow risk, and are ambiguity-averse. In the static setting, the first-best security is either convertible debt or levered equity. The optimal second-best security has an equity-like component in high cash flow states. Finally, if the two parties can renegotiate the contract after acquiring more information, the initial contract is risky debt. It is later renegotiated to a security with an equity component, and the conversion factor depends on the information acquired in the interim.

---

\*We are grateful to Archishman Chakraborty, David Dicks, Simon Gervais, Itzhak Gilboa, Itay Goldstein, William Mann, Andrey Malenko, Dmitry Orlov, Amiyatosh Purnanandam, Alejandro Rivera, Francesco Sangiorgi, Martin Szydlowski, Luke Taylor, Giulio Trigilia, Pavel Zryumov, and participants at the Cavalcade, EFA, IDC Israel, Minnesota Corporate Finance, Northern Finance Association, and WFA-CFAR Corporate Finance conferences and seminars at Hawaii and Rochester for helpful comments.

<sup>†</sup>George Mason University; [slee65@gmu.edu](mailto:slee65@gmu.edu)

<sup>‡</sup>Stephen M. Ross School of Business, University of Michigan; [urajan@umich.edu](mailto: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.<sup>1</sup> Indeed, according to Knight, profit is a reward to an entrepreneur for bearing uncertainty.

This uncertainty is increasingly being recognized in angel financing. For example, the recent Angel Funders Report released by the Angel Capital Association in the U.S. documents that in their sample 36% of angel financing rounds consisted of convertible notes or SAFEs (Simple Agreements for Future Equity).<sup>2</sup> Both these kinds of securities offer the investor conversion to equity when a future qualifying event (either a new financing round or an exit) occurs. Importantly, the price at which the conversion will occur is not specified *ex ante*, but instead will depend on the price set in the future when the qualifying event occurs. The standard reason provided for this is that the valuation is too uncertain at the time of the angel financing round.<sup>3</sup>

It is therefore important to understand the effect of uncertainty on the financing of small startup firms. As the Ellsberg (1961) paradox demonstrates, individuals are averse to uncertainty, preferring gambles with known probabilities to those with unknown probabilities. In this paper, we consider the implications of ambiguity-aversion on the part of both an entrepreneur and investors for security design.

---

<sup>1</sup>In contrast, outcomes with known probabilities are termed “risky” rather than “uncertain.”

<sup>2</sup>See <https://www.prnewswire.com/news-releases/angel-capital-association-study-provides-new-insights-on-angel-gro.html>.

<sup>3</sup>See, for example, the definition of “Convertible Note” provided by the *Financial Times*, at <http://lexicon.ft.com/Term?term=convertible-note>.

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 have limited liability and 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. However, both are ambiguity-averse as in Hansen and Sargent (2001); that is, both parties are concerned about model misspecification, and are averse to this prospect.

We show that, in our setting, the value functions of the agents are equivalent to CARA (constant absolute risk aversion) utility functions. Thus, in the static setting, the optimal security involves sharing cash flow proportionally between the investors and the entrepreneur in high cash flow states. 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, which requires the investors to obtain more cash in low cash flow states. The division of cash flow in high cash flow states depends on the degree of ambiguity aversion of both parties. If the entrepreneur has relatively high ambiguity-aversion, the security has an equity-like feature, but is non-linear.<sup>4</sup> If the entrepreneur has relatively low ambiguity-aversion, 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.

We then turn to a dynamic setting in which, after the initial security has been issued, the parties acquire more information about the future cash flows of the firm. The amount of uncertainty faced by the contracting parties therefore decreases. In the Hansen and Sargent (2001) approach, the latter is equivalent to a reduction in the degree of ambiguity aversion for each party. As in Hermalin and Katz (1991) and Dewatripont et al. (2003), we further assume that the investor observes the chosen level of effort. Then, the crucial aspect of renegotiation is that it separates the incentive and insurance problems. We show that the initial contract is risky debt, and the contract is later renegotiated to an efficient ambiguity-sharing contract. Importantly, in our framework, this latter contract is *linear*. Putting the two stages together, therefore, the overall contract may be thought of as convertible debt, with the conversion price dependent on the information acquired before

---

<sup>4</sup>Convertible securities in the venture capital setting often have complicated liquidation preferences that can create non-linearities in their payments; see Kaplan and Strömberg (2003).

renegotiation. Empirically, in Series B and later rounds of venture financing, the contract at that stage is often simple equity.<sup>5</sup>

There are, of course, many models that lead to convertible debt being the optimal contract for start-up financing, including Cornelli and Yosha (2003) (window-dressing and stage financing), Schmidt (2003) and Hellmann (2006) (double moral hazard) and Dewatripont et al. (2003) (risk-averse parties and renegotiation), Ozerturk (2008), (risk-shifting), and Ortner and Schmalz (2018) (heterogeneous beliefs). Our contribution comes from showing that Knightian uncertainty, an important friction faced by start-up firms and investors, also implies convertible debt, with the conversion factor being contingent on how the firm has fared between inception and the time that conversion occurs.

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 adopt 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 shadow costs for information. This motivates us to consider renegotiation of the financial contract when ambiguity aversion reduces over time.

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

---

<sup>5</sup>See, for example, <https://www.wsgr.com/publications/PDFSearch/entreport/Q42017/private-company-financing-trends.htm>, which shows that the use of non-participating equity in such rounds is widely prevalent and has increased over time.

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.<sup>6</sup>

Bewley (1989) builds a theory of innovation and entrepreneurship based on uncertainty aversion. In his setting, outcome 1 is preferred to outcome 2 if it is superior for every probability distribution in the set of priors. He shows that 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, Malenko and Tsoy (2018) 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. Importantly, the issuer is privately informed about its prospects in their model, so that the security serves as a signal. In contrast, we work with a model in which information is symmetric but there can be moral hazard on the part of the entrepreneur.

Work on moral hazard with Knightian uncertainty includes Lopomo et al. (2011), who work in the Bewley setting, and Izhakian and Zender (2017), who work in a framework in which perceived probabilities depend on uncertainty. Both these papers find that there can be flat portions in the compensation contract for the agent. Kellner (2017) shows that, with smooth ambiguity aversion, the optimal contract need not be monotone. 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

---

<sup>6</sup>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.

under-compensation following a low performance.

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. (2017) 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 3.2, with and without a monotonicity requirement on security payments. The solution to the optimal contracting with ex-post renegotiation is presented in Section 4, and a few features of our model are discussed in Section 5. Section 6 concludes.

## 2 Multiplier Preferences and Constraint Preferences

In this section, we briefly review multiplier preferences and constraint 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 a Bernoulli utility function  $u : Z \mapsto \mathbb{R}$ . Consider a sigma-algebra on  $X$ , and let  $\Delta(X)$  denote the set of all countably-additive probability measures on  $X$ . Then, given a probability measure  $q \in \Delta(X)$ , 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. With a slight abuse of notation, let  $\Delta(q)$  denote the set of probability measures equivalent to  $q$  (that is, the set of measures that agree with  $q$  on measure zero events).

Given any measure  $p \in \Delta(q)$ , the relative entropy  $R(p||q)$  is defined by

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

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 considered more costly to choose. Specifically, the decision-maker evaluates a payoff profile  $r$  with reference measure  $q$  according to

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

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,  $\theta$  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. As  $\theta$  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  $\theta$  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  $\theta \rightarrow \infty$ , 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  $\theta \rightarrow 0$ , the decision-maker becomes infinitely ambiguity averse.

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

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

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

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

Going forward, for the rest of the paper, we will assume that all parties have ambiguity-averse preferences represented as in equation 4. We also assume that agents are neutral toward risk (i.e., toward stochastic events with known probability distributions), so that  $u(y) = y$ . With risk-neutrality and ambiguity-aversion, the framework recovers a functional form similar to that of constant absolute risk aversion (CARA) utility to represent agents' preferences.<sup>7</sup>

We emphasize that the interpretation of  $\theta$  is very different under ambiguity aversion, as compared to under risk aversion. Hansen and Sargent (2001), in Chapter 6, show that multiplier preferences are closely related to constraint preferences, which may be written as:

$$\hat{V}(r) = \min_{p \in \Delta(q)} \int_X u(r(x)) dp(x), \text{ subject to } R(p||q) \leq \eta. \quad (5)$$

The right-hand side of the relative entropy constraint,  $\eta$ , depends on the amount of information available to the decision-maker. If the decision-maker has relatively precise information,  $\eta$  is low, so that he must evaluate the payoff profile  $r$  using a measure  $p$  that is close to  $q$ . Conversely, if the decision-maker has noisy information,  $\eta$  is large, and  $p$  may be far from  $q$ .

Now, the parameter  $\theta$  introduced in the earlier formulation is interpretable as the shadow cost or Lagrange multiplier on the relative entropy constraint in 5. As we show formally in Appendix A.1, the right-hand side of this constraint (i.e.,  $\eta$ ), is inversely related to the Lagrange multiplier  $\theta$ . For example, as  $\eta \rightarrow 0$  (or, as noted earlier, when  $\theta \rightarrow \infty$ ), in the limit maximizing  $\hat{V}(r)$  is equivalent to maximizing expected utility.

Consider a firm evolving through time. Over time, more information about the firm's future prospects becomes available. Suppose first that the all agents have a single prior distribution over the firm's cash flows, and update this prior distribution using Bayes' rule. Suppose also

---

<sup>7</sup>Maccheroni et al. (2006) show that multiplier preferences are a special case of a larger set of preferences, variational preferences, which are represented more as  $V(r) = \min_{p \in \Delta(q)} \int_X u(r(x)) dp(x) + \theta c(p, q)$ , where  $c(\cdot, q) : \Delta(q) \rightarrow [0, \infty]$  is a convex function on  $\Delta(q)$ . This general representation reduces to the exponential form in equation 4 if  $c(\cdot, q) = R(\cdot||q)$ ; that is, if relative entropy  $R(p||q)$  represents the cost of choosing  $p$  rather than  $q$  to evaluate the relevant payoff profile.



that the reference distribution  $q(x)$  represents the true distribution over cash flows. Then, under reasonable conditions, the posterior distribution will be closer to the true distribution, compared to the prior distribution.<sup>8</sup> More precisely, the relative entropy between the posterior distribution and the true distribution is lower than the relative entropy between the prior distribution and the true distribution.

In a setting with multiple priors, we can consider the agents as updating each distribution in the set of feasible priors as new information arrives. Then, for each element in the set of feasible priors, the relative entropy between the posterior and true distributions is smaller than the relative entropy between the prior and true distributions. In the constraint preference model, this would effectively amount to  $\eta$  decreasing, even though the underlying preferences over ambiguity remain unchanged. A reduction in  $\eta$  in turn increases the shadow price of the informational constraint. That is, with variational preferences, it would translate to the parameter  $\theta$  increasing over time. In contrast, if parties were risk-averse but ambiguity-neutral, and  $\theta$  is interpreted as the coefficient of absolute risk aversion, it would remain unchanged if preferences are invariant over time. That is, unlike in the case of risk aversion, interpreting  $\theta$  as a parameter related to ambiguity aversion implies that it decreases through time. We build upon this idea in Section 4.

### 3 Static Security Design Problem

We now consider the static 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 amount  $I$  must be raised from external investors. The project generates a cash flow  $x \in X = [0, \bar{x}]$  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

---

<sup>8</sup>In the limit, the law of large numbers applies — given a large number of i.i.d. signals, the posterior distribution converges to true one. See, for example, Cover and Thomas (2006), Theorem 11.2.1.

in  $a$ , so that  $\psi'(a) > 0$  and  $\psi''(a) > 0$ . In addition, we assume that  $\psi'(0) = \psi(0) = 0$ . Investors and entrepreneurs agree on the effect that the action  $a$  has on the reference measure induced over the cash flows. In particular, they believe that action  $a$  likely leads to a distribution  $F(x | a)$  over cash flows at date 1, with associated density  $f(x | a)$ . We assume that  $F(x | a)$  has full support over  $X$  and has no mass points. We further assume that  $f(x | a)$  satisfies the Monotone Likelihood Ratio Property (MLRP); that is,  $\frac{\partial}{\partial a} \left( \frac{f_a(x|a)}{f(x|a)} \right) > 0$ , where  $f_a(x | a) = \frac{\partial f(x|a)}{\partial a}$ .

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  $\theta_I$ . That is, the value to investors of a security  $r(x)$  when the action is  $a$  is given by

$$V_I(r, a) = -\theta_I \ln \left( \int_X e^{-\frac{r(x)}{\theta_I}} f(x | a) dx \right) \quad (6)$$

Similarly, the entrepreneur evaluates risky cash flows according to equation 4, with ambiguity-aversion parameter  $\theta_E$ . 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

$$V_E(r, a) = -\theta_E \ln \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) - \psi(a). \quad (7)$$

Recall that  $\theta_I$  and  $\theta_E$  are preference parameters in this specification. As is standard in moral hazard settings, we assume that the preferences of the two parties with respect to the cash flows (and hence these parameters) are not affected by the choice of action  $a$ . That is,  $a$  affects the reference distribution over cash flows, but not the model uncertainty perceived by the investors and the entrepreneur. We assume that investors and entrepreneur have the same reference density  $f(x | a)$  in mind.<sup>9</sup>

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. As  $-\theta_I \ln \left( \int_X e^{-\frac{I}{\theta_I}} f(x | a) dx \right) = I$ , we can write the investors' individual rationality (IR) constraint as:

$$-\theta_I \ln \left( \int_X e^{-\frac{r(x)}{\theta_I}} f(x | a) dx \right) \geq I. \quad (8)$$

---

<sup>9</sup>Alternatively, if investors evaluate cash flows based on a reference density  $f_I(x | a) \neq f(x | a)$ , the relative optimism or pessimism of each party will depend both on the coefficient  $\theta$  and their own reference measure. As a result, the optimal security will depend at any  $x$  on both  $f_I(x | a)$  and  $f(x | a)$ .

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

$$a = \arg \max_{\tilde{a}} -\theta_E \ln \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | \tilde{a}) dx \right) - \psi(\tilde{a}). \quad (9)$$

For now, we assume the first-order approach is valid (in Section 3.4, we specify a sufficient condition for this). We therefore replace the IC condition in equation 9 with the corresponding first-order condition

$$-\theta_E \frac{\int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x | a) dx}{\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx} - \psi'(a) = 0. \quad (10)$$

Finally, we assume that the entrepreneur's reservation utility is zero. The complete contracting problem may therefore be stated as:

$$[\text{Problem P1}] \quad \max_{r(x), a} \quad -\theta_E \ln \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) - \psi(a) \quad (11)$$

$$\text{subject to:} \quad (\text{IR}) \quad -\theta_I \ln \left( \int_X e^{-\frac{r(x)}{\theta_I}} f(x | a) dx \right) \geq I \quad (12)$$

$$(\text{IC}) \quad -\theta_E \frac{\int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x|a) dx}{\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x|a) dx} - \psi'(a) = 0. \quad (13)$$

$$(\text{LL}) \quad 0 \leq r(x) \leq x \text{ for all } x. \quad (14)$$

Here, (LL) represents the limited liability constraints on the security.

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.

$$[\text{Problem P2}] \quad \max_{r(x), a} \quad -e^{\frac{\psi(a)}{\theta_E}} \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) \quad (15)$$

$$\text{subject to} (\text{IR2}) \quad -\int_X e^{-\frac{r(x)}{\theta_I}} f(x | a) dx \geq -e^{-\frac{I}{\theta_I}} \quad (16)$$

$$(\text{IC2}) \quad \int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x | a) dx + \frac{\psi'(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx = 0 \quad (17)$$

$$(\text{LL}) \quad 0 \leq r(x) \leq x \text{ for all } x. \quad (18)$$

Observe that, as mentioned earlier, the transformed functions for the entrepreneur and investor in 15 and 16 are identical to utility functions that represent constant absolute risk aversion (CARA) preferences.

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

All proofs are contained in Appendix A.2.

For the rest of the paper, we assume that the constraint set for problem P2 is non-empty. Given the conditional density over cash flows,  $f(x | a)$ , the cost of effort  $\psi(a)$ , and the investor's ambiguity-aversion parameter  $\theta_I$ , we essentially require that the required investment level  $I$  is sufficiently low. For example, suppose that  $\int_X e^{-\frac{x}{\theta_I}} f(x | 0) dx \leq e^{-\frac{I}{\theta_I}}$ . Then, the security and effort pair  $(r(x) = x, a = 0)$  is in the feasible set.

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  $\theta_I \rightarrow \infty$  and  $\theta_E \rightarrow \infty$ , 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.<sup>10</sup> (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 any action  $\tilde{a} \neq a$ , investors can give the entrepreneur zero cash and retain the entire output  $x$  for themselves.

---

<sup>10</sup>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.

Let  $\lambda$  denote the shadow price for the investors' IR constraint in equation 16. Further, for each  $x$ , let  $\underline{\gamma}_x$  denote the shadow price on the constraint  $r(x) \geq 0$  and  $\bar{\gamma}_x$  the shadow price for the constraint  $r(x) \leq x$ . Then, the Lagrangian for the first-best problem may be written as:

$$\begin{aligned} \mathcal{L}_f(r, a, \lambda) = & -e^{\frac{\psi(a)}{\theta_E}} \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) - \lambda \left[ \int_X e^{-\frac{r(x)}{\theta_I}} f(x | a) dx - e^{-\frac{I}{\theta_I}} \right] \\ & - \int_X \left[ -\underline{\gamma}_x r(x) + \bar{\gamma}_x (r(x) - x) \right] dx \end{aligned} \quad (19)$$

Let  $a_f$  denote the optimal effort level in the first-best problem,  $r_f$  the optimal security, and  $\lambda_f$  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  $f(x | a)$ , 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*

$$r_f(x) = \min \left\{ x, \left( \frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I \theta_E}{\theta_I + \theta_E} \left( \ln \frac{\lambda_f \theta_E}{\theta_I} - \ln e^{\frac{\psi(a_f)}{\theta_E}} \right) \right)^+ \right\}. \quad (20)$$

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 20 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 the project cash flow.<sup>11</sup> The linear term  $\frac{\theta_I}{\theta_I + \theta_E} x$  reflects optimal ambiguity-sharing between investors

---

<sup>11</sup>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.

and entrepreneur. The linearity of the term follows from the exponential form of the expressions in problem P2, the transformed problem.<sup>12</sup> It follows that the payoff on the security is weakly increasing in  $x$ , and is overall piecewise linear in  $x$ .

Recall that  $\theta_I$  is inversely related to the degree of ambiguity aversion expressed by the investors, and likewise  $\theta_E$  is inversely related to the ambiguity aversion of the entrepreneur. The more ambiguity-averse an agent is (i.e., the lower  $\theta$  is), the further (and so the more pessimistic) the distribution under which they evaluate the cash flows is, compared to the reference measure  $f(x | a)$ . 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, suppose that, keeping  $\theta_E$  fixed,  $\theta_I$  increases; i.e., the investors become less ambiguity-averse relative to the entrepreneur. As a result, in the new first-best contract, investors get relatively more cash in the high cash flow states. Of course, as  $\theta_I$  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.

The optimal action  $a_f$  satisfies the first-order condition  $\frac{\partial \mathcal{L}_f}{\partial a} = 0$ , or

$$e^{\frac{\psi(a)}{\theta_E}} \left( \frac{\psi'(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx + \int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x | a) dx \right) + \lambda \int_X e^{-\frac{r(x)}{\theta_I}} f_a(x | a) dx = 0. \quad (21)$$

The three equations, 16, 20, and 21 can be used to solve for  $a_f$ ,  $\lambda_f$ , and  $r_f$ . To illustrate the properties and comparative statics of the first-best contract, we consider the following numeric example.

### Example 1: First-best outcome

Let  $X = [0, 1]$ , and let the action set be  $A = [0, 1]$ . Set  $f(x | a) = 1 + a(2x - 1)$ , so that  $f_a(x | a) = 2x - 1$  and  $f_{aa}(x | a) = 0$ . 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  $I = 0.3$ .

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  $\int_X x f_a(x | a) dx = \psi'(a)$ , which in this example yields  $a_N^* = \frac{1}{6}$  (the subscript  $N$  denotes that both parties are ambiguity-neutral). Any

---

<sup>12</sup>Recall that, as shown by Wilson (1968), optimal risk-sharing with exponential utilities entails a linear sharing rule.

division of the cash flows such that in expectation the investors obtain  $I$  and the entrepreneur obtains  $\int_X xf(x | a) - I$  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  $\theta_E$  fixed at 1, and varying  $\theta_I$ . In each case, the security issued in the first-best case includes a substantial equity component.

1. Convertible debt or convertible preferred equity.

This security emerges if  $\theta_I$  is sufficiently low, relative to  $\theta_E$ . 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 or preferred equity in the low states.<sup>13</sup> 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 or preferred equity, with the conversion threshold set equal to the face value of the initial security.

2. Levered Equity.

This security emerges if  $\theta_I$  is sufficiently high, relative to  $\theta_E$ . 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.

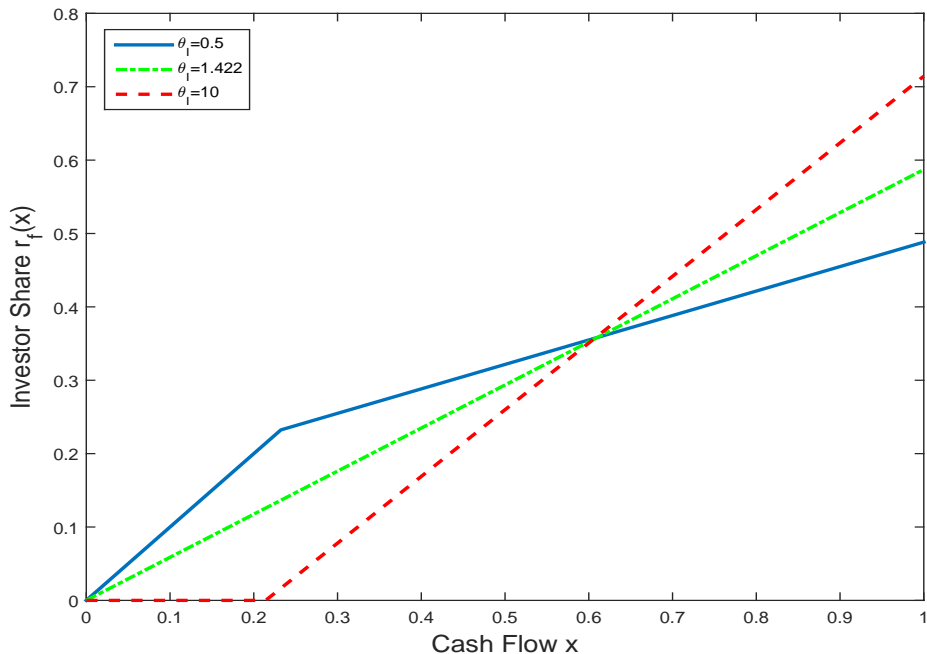
This is a knife-edge case that emerges at a specific value of  $\theta_I$ ; in the example, at  $\theta_I$  approximately equal to 1.422.

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

---

<sup>13</sup>In our setting, the security pays out only at time 1. Thus, control rights are not an issue, and so a security that has all rights to cash flows in low states may be thought of as either debt or preferred equity.

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  $\theta_I$  varies in Figure 1.



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

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

The optimal effort in the first-best contract falls as  $\theta_I$  increases, which is intuitive. An increase in  $\theta_I$  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).

### 3.2 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 17, the Lagrangian for problem P2 may be written as:

$$\begin{aligned}
\mathcal{L}(r, a, \lambda) = & -e^{\frac{\psi(a)}{\theta_E}} \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) - \lambda \left[ \int_X e^{-\frac{r(x)}{\theta_I}} f(x | a) dx - e^{-\frac{I}{\theta_I}} \right] \\
& + \mu \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x | a) dx + \frac{\psi'(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) \\
& - \int_X \left[ -\underline{\gamma}_x r(x) + \bar{\gamma}_x (r(x) - x) \right] dx. \tag{22}
\end{aligned}$$

Here,  $\mu$  is the shadow price on the entrepreneur's IC constraint, and, as before,  $\lambda$  is the shadow price on the investors' IR constraint,  $\underline{\gamma}_x$  the shadow price on the constraint  $r(x) \geq 0$ , and  $\bar{\gamma}_x$  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 (and a corresponding increase in the cash flow they obtain in low 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*

$$r^*(x) = \min \left\{ x, \left( \frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I \theta_E}{\theta_I + \theta_E} \left\{ \ln \frac{\lambda^* \theta_E}{\theta_I} - \ln \left( e^{\frac{\psi(a^*)}{\theta_E}} + \mu^* \left( \frac{f_a(x | a^*)}{f(x | a^*)} + \frac{\psi'(a^*)}{\theta_E} \right) \right) \right\} \right)^+ \right\}. \tag{23}$$

For any value of  $x$  at which both investors and entrepreneur obtain some of the cash flow, equation 23 implies that  $r^*(x) = \frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I \theta_E}{\theta_I + \theta_E} \left\{ \ln \frac{\lambda^* \theta_E}{\theta_I} - \ln \left( e^{\frac{\psi(a^*)}{\theta_E}} + \mu^* \left( \frac{f_a(x | a^*)}{f(x | a^*)} + \frac{\psi'(a^*)}{\theta_E} \right) \right) \right\}$ . The term  $\frac{\theta_I}{\theta_I + \theta_E} x$  is familiar from the first-best contract, and 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 in Innes (1990), incentive compatibility may not bind in this case. Recall that in the Innes model, 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 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 second-best case, the first-order condition in  $a$  is  $\frac{\partial \mathcal{L}}{\partial a} = 0$ , which reduces to

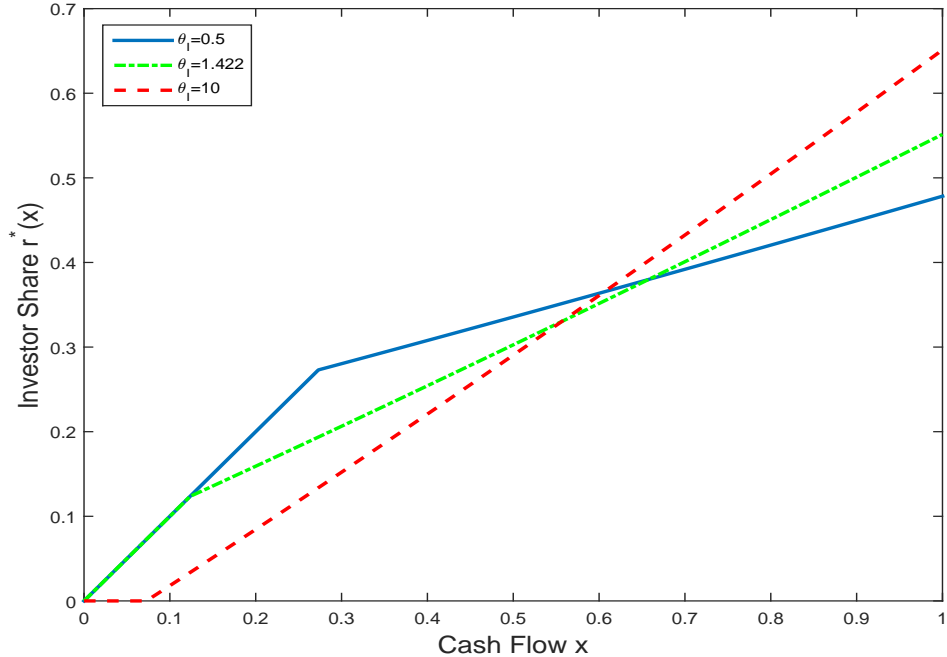
$$\lambda \int_X e^{-\frac{r(x)}{\theta_I}} f_a(x | a) dx - \mu \left\{ \frac{\psi''(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx + \frac{\psi'(a)}{\theta_E} \int_X e^{-\frac{x-r(x)}{\theta_E}} f_a(x | a) dx + \int_X e^{-\frac{x-r(x)}{\theta_E}} f_{aa}(x | a) dx \right\} = 0. \quad (24)$$

The four conditions represented by the above equation, the investors’ IR condition 16, the entrepreneur’s IC condition 17, and the equation for the optimal contract 23 can be used to solve for  $a^*$ ,  $\lambda^*$ ,  $\mu^*$ , and  $r^*$  in the case that the IC condition binds.

### Example 2: Second-best outcome

We use the same parameters as in Example 1. We set  $\theta_E = 1$ ,  $I = 0.3$ ,  $f(x | a) = 1 + a(2x - 1)$ , and  $\psi(a) = \frac{1}{2}a^2$ . The first-best security for three different values of  $\theta_I$  is exhibited in Figure 1. We illustrate the optimal security in the second-best setting in Figure 2. To highlight 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  $\theta_I$ . In the figures, the contrast is greatest for the intermediate  $\theta_I$  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 (although it is hard to see in the figure) the securities in Figure 2 are *not* piecewise linear—for high cash flows, there is a slight non-linearity in the security payoffs.



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

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

Therefore, they cannot be thought of directly in terms of equity. Nevertheless, for these parameter values, the securities have a component that to a large degree resembles equity.

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 Table 1.

Consider the comparative statics of the optimal security as  $\theta_E$  changes. We use the same cost function, conditional cash flow density, and investment level as before. We set  $\theta_I = 4$  and vary  $\theta_E$

$\theta_I$	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

across three levels, 1, 4, and 20. The optimal security in each case is exhibited in Figure 3.

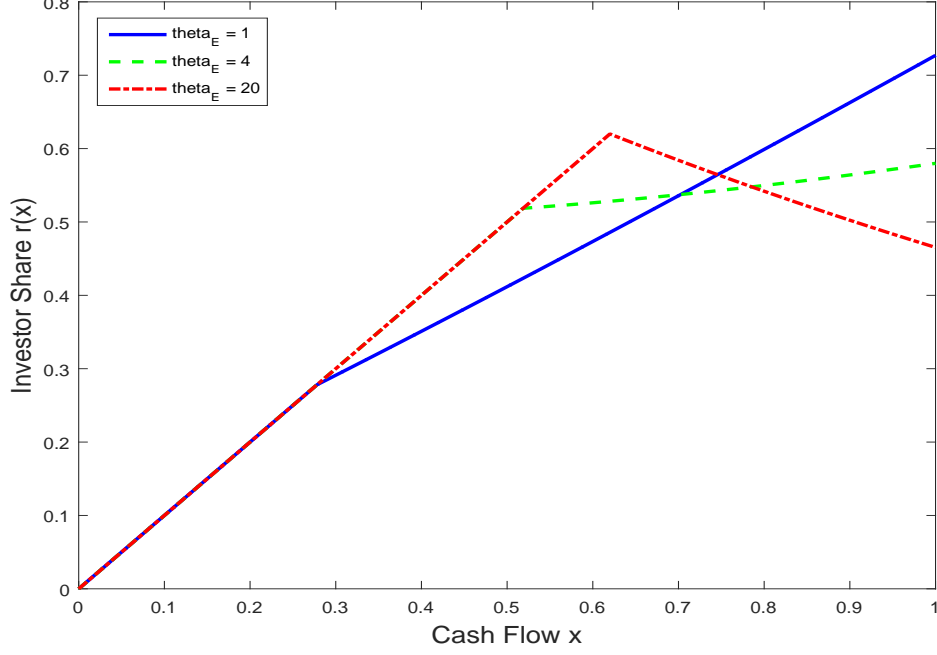
The broad intuition is as follows. The optimal security balances out the need for ambiguity-sharing, which entails giving more cash to the less ambiguity-averse party in high states, with the need to provide incentives to the entrepreneur, which entails giving more (often all) cash to the investor in low states. As  $\theta_E$  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.

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  $f(x | a)$ . In our example, when  $\theta_E = 20$ , the security payoffs decrease in the project cash flow when  $x$  exceeds approximately 0.6.

### 3.3 Second-Best Contract with Monotone Security Payoffs

As shown in Figure 3, when  $\theta_E$  is high relative to  $\theta_I$ , the optimal second-best security may provide a decreasing payoff to the investors when the cash flow increases. In such situations, we consider the implications of introducing a monotonicity restriction on the security, to ensure that the respective payoffs are non-decreasing in the cash flow. Following Innes (1990), 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  $x$ .*



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

Figure 3: Optimal Security as  $\theta_E$  Changes

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:

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

We build up to the contract that solves Problem P3 through the following steps. First, given Proposition 2, for some fixed values of  $\lambda, \mu, a$ , denote

$$\tilde{r}(x) = \frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I \theta_E}{\theta_I + \theta_E} \left\{ \ln \frac{\lambda \theta_E}{\theta_I} - \ln \left( e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x|a)}{f(x|a)} + \frac{\psi'(a)}{\theta_E} \right) \right) \right\}. \quad (25)$$

Next, 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.

Finally, consider monotonicity. Starting with  $\hat{r}(x)$ , we construct a monotonic security as follows. Set  $y_s = 0$ . At Step 1, define  $y_e = \sup\{x \geq y_s \mid \hat{r}'(x) \in [0, 1] \text{ for all } z \in [y_s, x]\}$ . Let  $r^M(x) = \hat{r}(x)$  for all  $x \leq y_e$ . By definition of  $y_e$ , the security  $\hat{r}(x)$  is non-monotone for  $x$  just greater than  $y_e$ . Set  $y_s = y_e$ . Now, there are two possibilities:

Step 2a: There exists a range  $(y_s, y_s + \epsilon)$  such that, for all  $x$  in this range,  $\hat{r}'(x) < 0$ . In this case, define  $y_e = \inf\{x \geq y_s \mid \hat{r}(x) = \hat{r}(y_s) \text{ and } \hat{r}'(x) > 0\}$ . Define  $r^M(x) = r^M(y_s)$  for all  $x \in (y_s, y_e]$ , and set  $y_s = y_e$ .

Step 2b: There exists a range  $(y_s, y_s + \epsilon)$  such that, for all  $x$  in this range,  $\hat{r}'(x) > 1$ . In this case, define  $y_e = \inf\{x \geq y_s \mid \hat{r}(x) = \hat{r}(y_s) + x - y_s \text{ and } \hat{r}'(x) < 1\}$ . Define  $r^M(x) = r^M(y_s) + x - y_s$  for all  $x \in (y_s, y_e]$ , and set  $y_s = y_e$ .

If  $y_e = \bar{x}$ , we are done; else, repeat the process starting with Step 1.

The optimal security once condition (M) is added to the problem then has the form given by  $r^M$ . Let  $(a^M, \lambda^M, \mu^M)$  denote values at which problem P3 is solved.

**Proposition 3.** *In any solution to problem P3,*

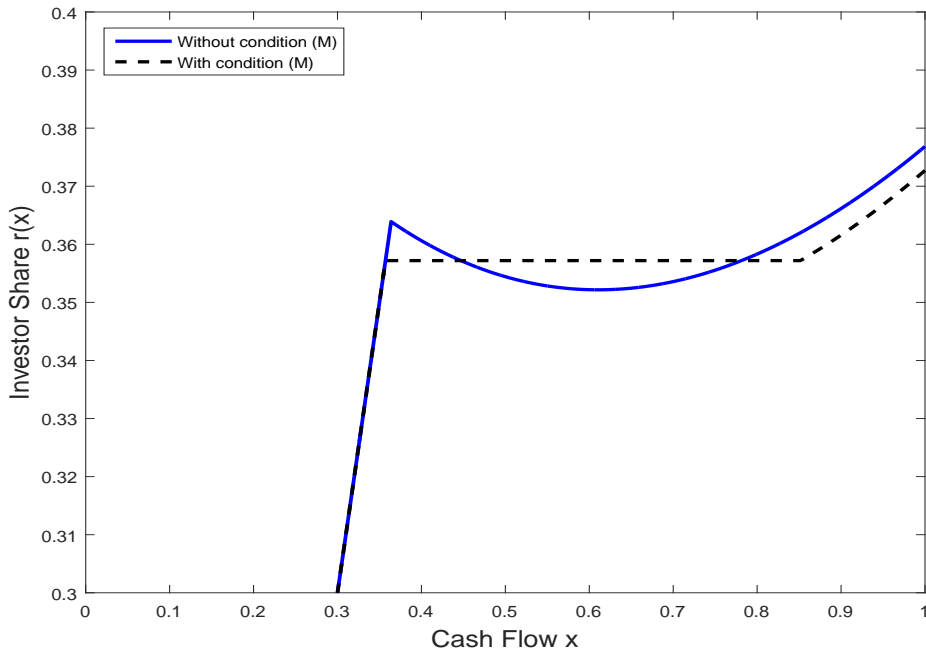
(i) *The investors' IR constraint binds.*

(ii) *The security issued to the investors is  $r^M(x)$ , evaluated at  $(a^M, \lambda^M, \mu^M)$ .*

To illustrate the application of Proposition 3, we revert to our running example. Let  $\psi(a) = \frac{1}{2}a^2$  and  $f(x|a) = 1 + a(2x - 1)$ . 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  $x > 0.53$ .

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  $r^M(x)$ . The dashed black line indicates the optimal security when the monotonicity condition is imposed. Note that this security resembles convertible debt with a relatively high conversion threshold. Although not apparent from the figure, when the cash flow is between 0.85 and 1, the security is slightly non-linear.



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

Figure 4: Optimal Security With and Without Monotonicity Requirement

Overall, then, we find that when the entrepreneur is ambiguity-averse (i.e.,  $\theta_E$  is low) and investors are mildly ambiguity-averse relative to the entrepreneur (i.e.,  $\theta_I$  is sufficiently high relative to  $\theta_E$ ), the optimal security has strictly increasing payoffs. When the entrepreneur is mildly ambiguity-averse (i.e.,  $\theta_E$  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  $\theta_I$  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  $\theta_I$  is high).

### 3.4 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 7, let  $V_E$  denote the entrepreneur's payoff. Let  $w(x) = x - r(x)$ ; then  $V_E = -\theta_E \ln \left( \int_X e^{-\frac{w(x)}{\theta_E}} f(x | a) dx \right) - \psi(a) = -\theta_E \ln \left( e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{w(x)}{\theta_E}} f(x | a) dx \right)$ . Denote  $U = e^{-\frac{V_E}{\theta_E}} = e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{w(x)}{\theta_E}} f(x | a) dx$ . 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  $\frac{\partial^2 U}{\partial a^2} = e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{w(x)}{\theta_E}} f(x | a) \left[ \frac{\psi''(a)}{\theta_E} + \frac{\psi'(a)}{\theta_E} \frac{f_a(x | a)}{f(x | a)} + \frac{f_{aa}(x | a)}{f(x | a)} \right] dx$ .

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

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

If this 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

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

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

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

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  $f(x | a) = 1 + a(2x - 1)$ , so that  $f_a(x | a) = 2x - 1$  and  $f_{aa}(x | a) = 0$ . 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,

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

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.

## 4 Dynamic Security Design: Renegotiation

We now consider optimal contracts in a setting in which  $\theta$  increases over time as additional information about the firm is revealed. In particular, we assume that after the initial financing is obtained, both entrepreneur and investors obtain some information about the time 1 cash flow. As discussed on Page 8, additional information reduces the amount of uncertainty faced by the parties, which in the model translates to an increase in their respective ambiguity-aversion parameters. Any change in the respective preference parameters, in turn, generates gains to trade between the entrepreneur and the investor, which allows room for renegotiating the optimal contract at date 0.5.

In practice, at date 0, neither party can be certain about how their information set will evolve in the future. They are, however, aware that there is a chance that more information will be acquired once the firm is up and running. In this section, to illustrate the effect of additional information on both the initial and final contracts, we make the strong assumption that at date 0 both parties can predict with certainty their information set at date 0.5, and hence know how their preference parameters will change after the contract is signed. We discuss relaxing this assumption later, in Section 4.1.

Specifically, we assume that the ambiguity aversion parameter of the entrepreneur increases from  $\theta_{E0}$  to  $\theta_{E1}$  as a result of the new information. Similarly, the ambiguity aversion parameter of the investors increases from  $\theta_{I0}$  to  $\theta_{I1}$ . For simplicity, we assume that the reference probability measure  $q$  used by both parties remains the same as before.

To keep the renegotiation problem tractable, we assume, following Hermalin and Katz (1991) and Dewatripont et al. (2003), 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.<sup>14</sup>

---

<sup>14</sup>A more complicated approach to renegotiation is to assume that effort is not observed by the investors. As Fudenberg and Tirole (1990) show, this results in a menu of contracts being considered at the renegotiation stage. Matthews (2001) considers renegotiation with unobserved effort, the entrepreneur having all bargaining power, and menus assumed to be infeasible, and determines conditions under which the optimal contract at date 0 is risky debt, as in Dewatripont et al. (2003) and in our Proposition 4 below.

The fact that effort is sunk provides a second source of gains to trade at the renegotiation stage, as 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 between the entrepreneur and the investor. Figure 5 provides the sequence of events in the model.

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.

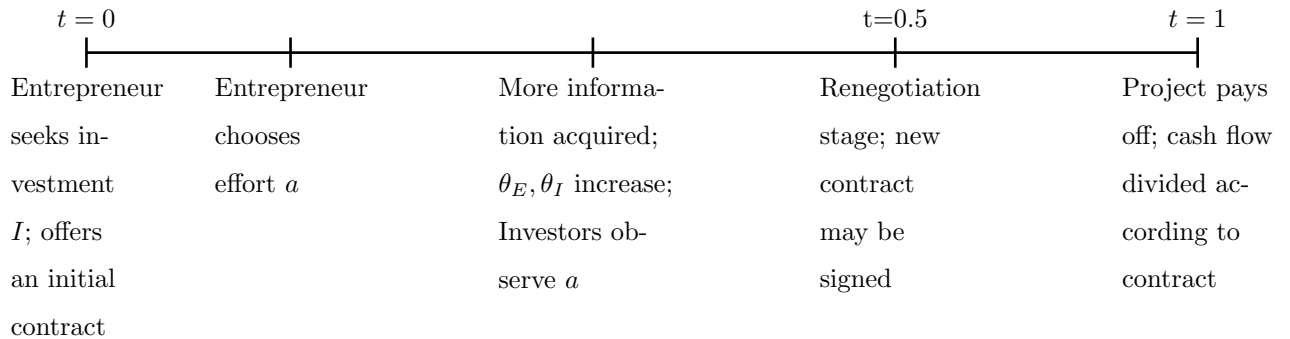


Figure 5: **Sequence of events with renegotiation**

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. As effort is no longer an issue at the renegotiation stage, the parties renegotiate the initial contract to an efficient ambiguity-sharing contract that is indeed linear.

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

$$\mathcal{C} := \{r(\cdot) \in C[0, \bar{x}] \mid 0 \leq r(x) \leq x \text{ for all } x, r'(x) \in [0, 1] \text{ for almost all } x\}. \quad (30)$$

Let  $r_0(\cdot)$  and  $r_n(\cdot)$  respectively be the initial and the renegotiated contracts that the entrepreneur offers to the investors. Each of these contracts belongs to the set  $\mathcal{C}$ . 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 the investors'

utility given the old contract  $r_0(\cdot)$ , effort  $a$  (which is known to the investor before renegotiation occurs), and their new ambiguity aversion parameter. That is,

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

Then,  $\hat{u}$  is the reservation utility of the investors at the renegotiation stage. At this stage, the entrepreneur offers the investors a contract that maximizes his own utility, subject to the constraint that investors are no worse off under the new contract as compared to the old contract. Recalling that we are working with a decreasing transformation of the utility function, and that the cost of effort is sunk at this stage, the transformed value functional for the entrepreneur at the renegotiation-stage is then:

$$\begin{aligned} T(r_0, a) &:= \max_{r_n(\cdot) \in \mathcal{C}} - \int_X e^{-\frac{x-r_n(x)}{\theta_{E1}}} f(x | a) dx \\ &\text{subject to} \quad - \int_X e^{-\frac{r_n(x)}{\theta_{I1}}} f(x | a) dx \geq -\hat{u}(r_0(\cdot), a; \theta_{I1}). \end{aligned} \quad (32)$$

Define  $\hat{r}_n(\cdot | r_0, a)$  to be the contract that achieves the minimum in equation 32.

At date 0, the entrepreneur chooses the optimal initial contract  $r_0(\cdot)$  and the optimal effort  $a$  to minimize the transformed value functional  $T(r_0, a)$ . Investors must, of course, obtain at least their reservation utility over the course of the game.

We assume that the entrepreneur and the investors are aware at the outset that their preference parameters will evolve from  $\theta_{j0}$  (where  $j \in \{E, I\}$ ) at date 0 to  $\theta_{j1}$  at date 1, and are dynamically consistent. Under this assumption, the time 0 preferences for either the entrepreneur or the investors are irrelevant to the contract. First, consider the entrepreneur. A dynamically consistent entrepreneur at date 0 evaluates uncertain cash flows at date 1 according to the preference parameter  $\theta_{E1}$ . Thus, the optimal effort given a contract at date 0 solves equation 11 with  $\theta_{E1}$  substituted in for  $\theta_E$ . Using the same transformation that results in equation 15, the objective function at time 0 may be written as  $\min_{r_0(\cdot) \in \mathcal{C}, a} e^{\frac{\psi(a)}{\theta_{E1}}} T(r_0, a)$ .

Investors, in turn are dynamically consistent. At date 0, they recognize that by the time the project yields cash, their ambiguity aversion parameter will have changed to  $\theta_{I1}$  and that, after renegotiation, they will obtain the security  $\hat{r}_n(\cdot)$ . It then follows that only the new optimal contract  $\hat{r}_n$  must satisfy individual rationality for investors, given  $\theta_{I1}$  and the effort  $a$  chosen at date 0.

We assume that the initial contract  $r_0$  must also satisfy limited liability and monotonicity. Thus,

the entrepreneur's problem at date 0 can be stated as:

$$\max_{r_0(\cdot) \in \mathcal{C}, a} e^{\frac{\psi(a)}{\theta_{E1}}} T(r_0, a) \quad (33)$$

$$\text{subject to : } - \int_X e^{-\frac{\hat{r}_n(x|r_0, a)}{\theta_{I1}}} f(x | a) dx \geq -e^{-\frac{I}{\theta_{I1}}}. \quad (34)$$

At this point, the security design problem above may be seen as a special case of the general problem analyzed in Dewatripont et al. (2003). The particular insight we offer is that in our framework with ambiguity aversion, the arrival of new information yields a setting that is mathematically equivalent to a change in the preference parameters of the parties. Further, given such a change, with dynamically consistent agents, only their new preferences affect both the renegotiated contracts and the contract at date 0.

As expected from Dewatripont et al. (2003), the optimal security at date 0 is risky debt, which is of course a piecewise-linear contract. We formally prove this in the next Proposition. As effort is sunk at the renegotiation stage, it is immediate that, in equilibrium, the IR condition 34 must bind. Among feasible contracts and associated efforts that satisfy this binding IR condition, we show that it is optimal for the entrepreneur to offer a risky debt as an initial contract. Therefore, 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 post-renegotiation value of the contract to the entrepreneur.

The intuition underlying the optimal contract is as follows. No matter what the initial contract is, the final contract after renegotiation will provide optimal ambiguity sharing between the investors and the entrepreneur. Thus, renegotiation effectively separates the incentive and the insurance problems that are associated with moral hazard. However, as long as the initial contract provides risky payoffs to the investors, the reservation utility of investors strictly increases with the entrepreneur's effort; that is,  $\partial \hat{u}(r_0, \cdot) / \partial a < 0$ . This increase in investor reservation utility acts as a wedge between the marginal benefit and the marginal cost of effort to the entrepreneur. The initial contract, therefore, must minimize this wedge. Among all date 0 securities, risky debt minimizes this effect.

**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  $\theta_E$  and  $\theta_{I1}$ . That is,  $r_n^*$  satisfies

$$r_n^*(x) = \min \left\{ x, \left( \frac{\theta_{I1}}{\theta_{I1} + \theta_{E1}} x + \frac{\theta_{I1}\theta_{E1}}{\theta_{I1} + \theta_{E1}} \left( \ln \frac{\lambda_n \theta_{E1}}{\theta_{I1}} - \ln e^{\frac{\psi(a^*)}{\theta_{E1}}} \right) \right)^+ \right\}, \quad (35)$$

where  $\lambda_n$  is the Lagrange multiplier associated with the inequality constraint in 32.

(ii) The entrepreneur's effort  $a^*$  is strictly lower than in the first-best problem given  $\theta_{E1}$  and  $\theta_{I1}$ .

As shown in Example 1, the security that solves the first-best problem is piecewise-linear, and may be any one of convertible debt, levered equity, or unlevered equity, depending on the exact parameter values.<sup>15</sup> 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 (or convertible preferred equity; see footnote 13) — it starts out as debt (or preferred equity), and either converts wholly or partially into equity at date 0.5, or at the minimum specifies a precise conversion option. In fact, Cornelli and Yosha (2003) show that the outcome with renegotiation can be replicated by issuing convertible debt at date 0. Importantly, in our setting, the conversion factor depends on the new ambiguity-aversion parameters of each party, and hence on the information acquired between times 0 and 0.5.

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.

---

<sup>15</sup>Note that the linearity of the first-best security in our setting emerges directly from the exponential form of the value functions for the respective agents. Other models of ambiguity-aversion will still lead to cash flows being shared between entrepreneur and investors in high states, but the sharing rule will in general be non-linear.

## 4.1 The more general problem

We now discuss relaxing our earlier assumption that the entrepreneur and investors know at date 0 how their information set will change by date 0.5. In general, we expect that at date 0, both parties are uncertain (in the Knightian sense) about the extent of information they will acquire in the future. In principle, this could be modeled with each party having a reference distribution in mind for the value of both  $\theta_E$  and  $\theta_I$  at time 0.5. When the renegotiation stage arrives, the parties know their respective information sets, which amounts to particular values of  $\theta_{E1}$  and  $\theta_{I1}$  being realized from the respective reference distributions.

The outcome at the renegotiation stage, therefore, will look exactly as above. Effort is sunk, so incentives are no longer an issue, and the renegotiation results in first-best ambiguity sharing between the two parties. The final contract between the two is piecewise-linear, and satisfies equation 35 in Proposition 4, given the realized values  $\theta_{E1}$  and  $\theta_{I1}$ .

To determine the optimal effort  $a^*$  and the date 0 contract  $r_0$ , we would need to explicitly recognize that  $T_0(\cdot)$  as defined in equation 32 depends on the realized values  $\theta_{E1}$  and  $\theta_{I1}$ . At date 0, Both entrepreneur and investor have a reference distribution in mind over these future parameter values. In other words, there is a double layer of uncertainty — they are uncertain over their information sets (and hence preference parameters) at date 0.5 as well as over the cash flows of the firm given a level of effort. Formally modeling this second layer of uncertainty over preference parameters, in the Hansen-Sargent framework, will require taking a negative exponential transformation of the value functions exhibited in equations 6 and 7 and taking an expectation with respect to the reference distribution over parameter values. Note that an additional uncertainty parameter is introduced to account for this second layer of uncertainty.

It is clear that optimal effort will now depend on the additional dimension of uncertainty, and on the entrepreneur's aversion to this uncertainty. Similarly, the investor's IR condition will hold in expectation, given the reference distribution over parameter values. In the main part of Section 4, we abstract away from these complications, and consider the simpler scenario in which evolution of information sets is deterministic, and so at date 0 the parties know their own future preferences. We leave the question of the optimal contract in the more complicated scenario as a project to be pursued later.

## 5 Discussion

As we have shown, the Hansen-Sargent approach to robustness combined with risk-neutrality recovers CARA preferences for each party. Two questions immediately arise. First, if risk aversion is enough to understand security design, why worry about Knightian uncertainty? Second, how can one distinguish between the effects of ambiguity aversion and risk aversion?

On the first question, we have argued in the Introduction that start-up firms face Knightian uncertainty. This can be seen both from the abstract notion that probabilities are hard to assign for new projects on which little data are available, and from the concrete finding that it is common in angel financing for investors to obtain convertible securities for which the valuation of the firm (and hence the conversion factor) is delayed until a future date, at which more information will be available. Thus, we argue that uncertainty is a plausible channel that leads to the forms of financing we see with start-up firms.

A related idea is that risk aversion with respect to a specific investment can be diversified away, so an investor who is risk-averse with respect to their total wealth should nevertheless be close to risk-neutral with respect to the outcome of any given investment. In contrast, as Uppal and Wang (2003) show, if different assets have different levels of uncertainty associated with their cash flows, investors choose to remain significantly under-diversified, and hence bear Knightian uncertainty. Thus, it is reasonable to model investors in start-up firms as ambiguity-averse.

On the second question, it is indeed correct that the outcomes under ambiguity-aversion and risk-aversion are very similar, with both giving rise to convertible securities at the initial date. In the Hansen-Sargent framework, we argue that if additional information is acquired over time, the effect is equivalent to the effect of ambiguity-aversion parameters changing through time. In other words, it is as if the agents' preferences are evolving through time. Thus the interpretation of the final contract is quite different than under risk aversion, where agents' preferences should be staying constant through time.

Suppose, instead, we consider a model with risk aversion for both parties (as in Dewatripont et al. (2003)). In such a model, new information between firm inception and contract renegotiation will lead to a refined posterior distribution over cash flows, with the risk aversion parameters remaining constant. It may be noted from equation 35 in Proposition 4 that the slope of the final contract depends only on the ratio  $\frac{\theta_{I1}}{\theta_{I1} + \theta_{E1}}$ ; that is, on the preference parameters of the parties

at the renegotiation stage. Therefore, our model with ambiguity-aversion predicts that the equity share obtained by the investor in high cash flow states will depend on the firm's evolution and on its prospects at the time the post-renegotiation security is issued. In contrast, a model with risk-averse parties would predict that this share is invariant to the firm's prospects at that stage. Anecdotal observation suggests that the latter prediction is easily falsified in favor of the former one.

## 6 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 in the static setting 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 ambiguity-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 with a conversion option in high cash flow states.

In a dynamic setting, over time more information about the firm becomes available. To the extent that multiplier preferences can also be represented as constraint preferences, it is reasonable to think of both parties becoming less ambiguity-averse in the future. We show that renegotiation in this setting eliminates non-linear components in the second-best contract, because incentive compatibility is no longer an 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, uncertainty is a friction that start-up firms face in practice, so our model provides a plausible explanation for the existence of such contracts.

## A Appendix

### A.1 Relationship between multiplier preferences and constraint preferences

Define the constraint preference problem

$$J(\eta) \triangleq \min_{p \in \Delta(q)} \left\{ \int u(x)p(x)dx : R(p||q) \leq \eta \right\} \quad (36)$$

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

$$\mathcal{L} = \int u(x)p(x)dx + \theta (R(p||q) - \eta) \quad (37)$$

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

$$L(\theta) \triangleq \min_{p \in \Delta(q)} \left\{ \int (u(x) - \theta \eta) p(x)dx + \theta R(p||q) \right\} \quad (38)$$

Assume that strong duality holds. Then, we have:

**Lemma A1.** *Suppose the minimum of  $J(\eta)$  is finite. Then there exists a  $\theta > 0$  such that:*

$$J(\eta) = \max_{\theta > 0} L(\theta) \quad (39)$$

*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  $\theta$  satisfies the first-order condition:

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

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

## A.2 Proofs

### Proof of Lemma 1

First, consider the IR constraint in equation 12. Dividing throughout by  $-\theta_I$ , we have

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

Taking the exponential of both sides yields the constraint (IR2) exhibited in equation 16.

Next, consider the IC constraint in equation 13. Multiply throughout by  $-\frac{\int_X e^{-\frac{x-r(x)}{\theta_E}} f(x|a)dx}{\theta_E}$  to obtain

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

which is the constraint (IC2) exhibited in equation 17.

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 11. Denote  $\Phi(r, a) = -\theta_E \ln \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) - \psi(a)$ . Then,

$$\begin{aligned} \Phi(r, a) &= -\theta_E \left[ \ln \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) + \frac{\psi(a)}{\theta_E} \right] \\ &= -\theta_E \left[ \ln \left( \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right) + \ln \left( e^{\frac{\psi(a)}{\theta_E}} \right) \right] \\ &= -\theta_E \ln \left( e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right). \end{aligned} \quad (40)$$

Now, maximizing  $\Phi(r, a)$  is equivalent to minimizing  $-\frac{1}{\theta_E} \Phi(r, a) = \ln \left( e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx \right)$ .

Finally, minimizing the last expression is equivalent to minimizing its exponential,

$e^{\frac{\psi(a)}{\theta_E}} \int_X e^{-\frac{x-r(x)}{\theta_E}} f(x | a) dx$ , or equivalently, maximizing the negative of the latter expression.

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

### 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  $\int_X e^{-\frac{\tilde{r}(x)}{\theta_I}} f(x | a) dx < e^{-\frac{I}{\theta_I}}$ . 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  $\int_X e^{-\frac{\hat{r}(x)}{\theta_I}} f(x | a) dx < e^{-\frac{I}{\theta_I}}$ . 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

$$\begin{aligned} \frac{e^{\frac{\psi(a)}{\theta_E}}}{\theta_E} e^{-\frac{x-r(x)}{\theta_E}} f(x | a) - \frac{\lambda}{\theta_I} e^{-\frac{r(x)}{\theta_I}} f(x | a) - \underline{\gamma}_x + \bar{\gamma}_x &= 0 \\ \left[ e^{-\frac{x-r(x)}{\theta_E} + \frac{\psi(a)}{\theta_E}} - \frac{\lambda \theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I}} \right] \frac{f(x | a)}{\theta_E} &= \underline{\gamma}_x - \bar{\gamma}_x. \end{aligned} \quad (41)$$

Now, there are three cases to consider.

Case 1:  $\underline{\gamma}_x > 0$ . Then,  $r(x) = 0$  by complementary slackness, so it follows that  $\bar{\gamma}_x = 0$ . Equation 41 reduces to

$$\left[ e^{-\frac{-x+\psi(a)}{\theta_E}} - \frac{\lambda \theta_E}{\theta_I} \right] \frac{f(x | a)}{\theta_E} = \underline{\gamma}_x. \quad (42)$$

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

$$x < \psi(a) - \theta_E \ln \left( \frac{\lambda \theta_E}{\theta_I} \right). \quad (43)$$

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:  $\bar{\gamma}_x > 0$ . Then,  $r(x) = x$  by complementary slackness, so it follows that  $\underline{\gamma}_x = 0$ . Equation 41 reduces to

$$\left[ e^{\frac{\psi(a)}{\theta_E}} - \frac{\lambda \theta_E}{\theta_I} e^{-\frac{x}{\theta_I}} \right] \frac{f(x | a)}{\theta_E} = -\bar{\gamma}_x. \quad (44)$$

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

$$x < \theta_I \left[ \ln \left( \frac{\lambda \theta_E}{\theta_I} \right) - \frac{\psi(a)}{\theta_E} \right]. \quad (45)$$

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:  $\bar{\gamma}_x = \underline{\gamma}_x = 0$ . Then,  $r(x) \in (0, x)$ . Here, equation 41 reduces to

$$\left[ e^{-\frac{x-r(x)}{\theta_E} + \frac{\psi(a)}{\theta_E}} - \frac{\lambda\theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I}} \right] \frac{f(x|a)}{\theta_E} = 0. \quad (46)$$

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

$$\begin{aligned} -\frac{x}{\theta_E} + \frac{r(x)}{\theta_E} + \frac{\psi(a)}{\theta_E} &= \ln\left(\frac{\lambda\theta_E}{\theta_I}\right) - \frac{r(x)}{\theta_I} \\ r(x) &= \frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I\theta_E}{\theta_I + \theta_E} \left( \ln\left(\frac{\lambda\theta_E}{\theta_I}\right) - \frac{\psi(a)}{\theta_E} \right). \end{aligned} \quad (47)$$

Now, let  $\lambda_f$  and  $a_f$  be the values of  $\lambda$  and  $a$  when the Lagrangian has been optimized. Denote the RHS of 47, 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 20, 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

$$\begin{aligned} \frac{e^{\frac{\psi(a)}{\theta_E}}}{\theta_E} e^{-\frac{x-r(x)}{\theta_E}} f(x|a) - \frac{\lambda}{\theta_I} e^{-\frac{r(x)}{\theta_I}} f(x|a) \\ + \frac{\mu}{\theta_E} \left( e^{-\frac{x-r(x)}{\theta_E}} f_a(x|a) + \frac{\psi'(a)}{\theta_E} e^{-\frac{x-r(x)}{\theta_E}} f(x|a) \right) - \underline{\gamma}_x + \bar{\gamma}_x &= 0 \\ e^{-\frac{x-r(x)}{\theta_E}} \left[ e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x|a)}{f(x|a)} + \frac{\psi'(a)}{\theta_E} \right) - \frac{\lambda\theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I} + \frac{x-r(x)}{\theta_E}} \right] \frac{f(x|a)}{\theta_E} &= \underline{\gamma}_x - \bar{\gamma}_x. \end{aligned} \quad (48)$$

Now, there are three cases to consider.

Case 1:  $\underline{\gamma}_x > 0$ . Then,  $r(x) = 0$  by complementary slackness, so it follows that  $\bar{\gamma}_x = 0$ . Equation 48 reduces to

$$e^{-\frac{x}{\theta_E}} \left[ e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x|a)}{f(x|a)} + \frac{\psi'(a)}{\theta_E} \right) - \frac{\lambda\theta_E}{\theta_I} e^{\frac{x}{\theta_E}} \right] \frac{f(x|a)}{\theta_E} = \underline{\gamma}_x. \quad (49)$$

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

$$e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) > \frac{\lambda\theta_E}{\theta_I} e^{\frac{x}{\theta_E}}. \quad (50)$$

Taking natural logs on both sides, we have

$$\ln \left( e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) \right) > \ln \frac{\lambda\theta_E}{\theta_I} + \frac{x}{\theta_E}. \quad (51)$$

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:  $\bar{\gamma}_x > 0$ . Then,  $r(x) = x$  by complementary slackness, so it follows that  $\underline{\gamma}_x = 0$ . Equation 48 reduces to

$$\left[ e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) - \frac{\lambda\theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I}} \right] \frac{f(x | a)}{\theta_E} = -\bar{\gamma}_x. \quad (52)$$

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

$$\begin{aligned} e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) &< \frac{\lambda\theta_E}{\theta_I} e^{-\frac{x}{\theta_I}} \\ \ln \left( e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) \right) &< \ln \frac{\lambda\theta_E}{\theta_I} - \frac{x}{\theta_I}. \end{aligned} \quad (53)$$

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:  $\underline{\gamma}_x = \bar{\gamma}_x = 0$ . Then,  $r(x) \in (0, x)$ . Here, equation 48 reduces to

$$e^{-\frac{x-r(x)}{\theta_E}} \left[ e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) - \frac{\lambda\theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I} + \frac{x-r(x)}{\theta_E}} \right] \frac{f(x | a)}{\theta_E} = 0, \quad (54)$$

which implies that

$$e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) = \frac{\lambda\theta_E}{\theta_I} e^{-\frac{r(x)}{\theta_I} + \frac{x-r(x)}{\theta_E}} \quad (55)$$

$$\ln \left( e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) \right) = \ln \frac{\lambda\theta_E}{\theta_I} - \frac{r(x)}{\theta_I} + \frac{x-r(x)}{\theta_E}. \quad (56)$$

The last equation directly implies that

$$r(x) = \frac{\theta_I}{\theta_I + \theta_E} x + \frac{\theta_I \theta_E}{\theta_I + \theta_E} \left\{ \ln \frac{\lambda\theta_E}{\theta_I} - \ln \left( e^{\frac{\psi(a)}{\theta_E}} + \mu \left( \frac{f_a(x | a)}{f(x | a)} + \frac{\psi'(a)}{\theta_E} \right) \right) \right\}. \quad (57)$$

Now, let  $a^*$ ,  $\lambda^*$ , and  $\mu^*$  denote the values of the respective variables when the Lagrangian has been optimized. Let  $\hat{r}(x)$  denote the RHS of equation 57. 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 57. As  $\lambda \rightarrow 0$ , regardless of the value of  $x$ , the term  $\ln \frac{\lambda \theta_E}{\theta_I} \rightarrow -\infty$ , 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

$$H(x, \lambda, \mu, r(\cdot), \rho(\cdot)) = -e^{\frac{\psi(a)}{\theta_E}} e^{-\frac{(x-r(x))}{\theta_E}} f(x|a) - \lambda \left( e^{-\frac{r(x)}{\theta_I}} f(x|a) - e^{-I/\theta_I} \right) + \mu \left( \frac{\psi'(a)}{\theta_E} e^{-\frac{(x-r(x))}{\theta_E}} f(x|a) + e^{-\frac{(x-r(x))}{\theta_E}} f_a(x|a) \right) + \xi(x)\rho(x), \quad (58)$$

where we temporarily suppress the limited liability constraints  $0 \leq r(x) \leq x$  and the constraint  $\rho(x) \leq 1$ . Further,  $\xi$  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)  $\rho(x) = \arg \max_{0 \leq \tilde{\rho}(x)} H(x, \lambda, \mu, \bar{r}(\cdot), \tilde{\rho}(\cdot)).$

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

$$\xi'(x) = -\frac{\partial H}{\partial r}(x). \quad (59)$$

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

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

$$\rho(x) = \arg \max_{0 \leq \rho(x) \leq 1} H(x, \lambda, \mu, r(\cdot), \rho(\cdot)) = \arg \max_{0 \leq \rho(x) \leq 1} \xi(x)\rho(x) \quad (60)$$

That is,

$$\rho(x) = \begin{cases} 0 & \text{if } \xi(x) > 0 \\ 1 & \text{if } \xi(x) < 0 \\ (0, 1) & \text{if } \xi(x) = 0 \end{cases} \quad (61)$$

The following cases emerge.

Case 1:  $\xi(x) = 0$  over a range of positive measure, so  $\xi'(x) = 0$  over this range. From equation 59, we have

$$\frac{\partial H}{\partial r}(x) = e^{-\frac{w(x)}{\theta_E}} f(x|a) \left( \frac{1}{\theta_E} e^{\frac{\psi(a)}{\theta_E}} - \frac{\lambda}{\theta_I} e^{-\frac{r(x)}{\theta_I} + \frac{x-r(x)}{\theta_E}} + \mu \left( \frac{\psi'(a)}{\theta_E^2} + \frac{1}{\theta_E} \frac{f_a(x|a)}{f(x|a)} \right) \right) = 0 \quad (62)$$

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 59, we have

$$e^{-\frac{(x-c_0)}{\theta_E}} f(x|a) \left( \frac{\lambda}{\theta_I} e^{\frac{x}{\theta_E} - \frac{\theta_E + \theta_I}{\theta_I \theta_E} c_0} - \mu \frac{1}{\theta_E} \frac{f_a(x|a)}{f(x|a)} - \frac{1}{\theta_E} e^{\frac{\psi(a)}{\theta_E}} - \mu \frac{\psi'(a)}{\theta_E^2} \right) = \xi'(x) \quad (63)$$

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 59, we have

$$e^{-\frac{\hat{c}_0}{\theta_E}} f(x|a) \left( \frac{\lambda}{\theta_I} e^{-\left(\frac{x}{\theta_I} - \frac{\theta_I + \theta_E}{\theta_E \theta_I} \hat{c}_0\right)} - \mu \frac{1}{\theta_E} \frac{f_a(x|a)}{f(x|a)} - \frac{1}{\theta_E} e^{\frac{\psi(a)}{\theta_E}} - \mu \frac{\psi'(a)}{\theta_E^2} \right) = \xi'(x) \quad (64)$$

Observe that the mapping

$$x \mapsto \frac{\lambda}{\theta_I} e^{-\left(\frac{x}{\theta_I} - \frac{\theta_I + \theta_E}{\theta_E \theta_I} \hat{c}_0\right)} - \mu \frac{1}{\theta_E} \frac{f_a(x|a)}{f(x|a)} - \frac{1}{\theta_E} e^{\frac{\psi(a)}{\theta_E}} - \mu \frac{\psi'(a)}{\theta_E^2} \quad (65)$$

is strictly decreasing in  $x$ .

Given the parameters  $(\theta_E, \theta_I)$ , the optimal values of  $(a, \lambda, \mu)$  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  $[0, \bar{x}]$ .



Otherwise,  $\xi(x) > 0$  for all  $x \in [0, \bar{x}]$ , which implies  $r(x) = 0$  for all  $x \in [0, \bar{x}]$ , 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  $x \in [\hat{x}_0, \bar{x}]$ . To see this, suppose that  $\xi(x) < 0$  for some  $\tilde{x} \in [\hat{x}_0, \bar{x}]$ . As the mapping in equation 65 is strictly decreasing, it must be that  $\xi'(x) < 0$  for all  $x > \tilde{x}$ . However, in this case, we have  $\xi(\bar{x}) < 0$ , which violates the boundary condition  $\xi(\bar{x}) = 0$ . Therefore,  $\xi(x) = 0$  on  $[\hat{x}_0, \bar{x}]$  and we have the situation in equation 62.

(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  $x \in [0, \bar{x}]$ , which would lead to a violation of the boundary condition  $\xi(\bar{x}) = 0$ . 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  $x \in [\hat{x}_0, \bar{x}]$ . Then,

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

(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 65 and the transversality condition  $\xi(\bar{x}) = 0$  ensure that  $\xi(x) = 0$  for  $x \in [\hat{x}_2, \bar{x}]$ . That is,  $r(x) = r^*(x)$  on  $[\hat{x}_2, \bar{x}]$ . Finally, if  $\hat{x}_2 = \bar{x}$ , then the security is standard debt.

The function  $r^M(x)$  as defined in the text encapsulates these various cases.

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

## Proof of Proposition 4

*Proof.* Once we show that a risky debt contract minimizes the marginal effect of effort to the investors' reservation utility  $|\partial \hat{u}(\cdot) / \partial a|$ , the proof is the same as in Proposition 2 in Dewatripont et al. (2003).

Suppose that  $r_0(\cdot)$  is an arbitrary initial contract in  $\mathcal{C}$ ; that is,  $r'(x) \in [0, 1]$  for all  $x$ . Then, the investors' reservation utility is increasing in effort:

$$\frac{\partial}{\partial a} \hat{u}(\cdot) = \int_X -e^{-\frac{r_0(x)}{\theta I_1}} f_a(x|a) dx = \int_X -e^{-\frac{r_0(x)}{\theta I_1}} \frac{f_a(x|a)}{f(x|a)} f(x|a) dx$$

$$= \mathbb{E}_{x \sim f(\cdot|a)} \left[ -e^{-\frac{r_0(x)}{\theta_{I1}}} \cdot \left( \frac{f_a(x|a)}{f(x|a)} \right) \right] \stackrel{(a)}{=} \mathbb{C}\text{OV} \left[ -e^{-\frac{r_0(x)}{\theta_{I1}}}, \left( \frac{f_a(x|a)}{f(x|a)} \right) \right] \stackrel{(b)}{\geq} 0. \quad (66)$$

Here, the equality (a) is valid because

$$\mathbb{E} \left[ \frac{f_a(x|a)}{f(x|a)} \right] = \int_X \frac{f_a(x)}{f(x)} f(x) dx = \int_X f_a(x) dx = \frac{\partial}{\partial a} \left( \int_X f(x) dx \right) = 0, \quad (67)$$

where the derivative and 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 mapping  $x \mapsto -e^{-\frac{r_0(x)}{\theta_{I1}}}$  is increasing by the monotonicity assumption on  $r_0(x)$ .

Now turn to show that a risky debt minimizes the magnitude  $|\partial \hat{u} / \partial a|$ . 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 reservation utility as  $\delta(\cdot)$ :

$$\int -e^{-\frac{r(x)}{\theta_{I1}}} f(x|a) dx = \int -e^{-\frac{\delta(x)}{\theta_{I1}}} f(x|a) dx. \quad (68)$$

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

$$\begin{aligned} \left| \frac{\partial}{\partial a} \hat{u}(r, a) \right| - \left| \frac{\partial}{\partial a} \hat{u}(\delta, a) \right| &= - \int_0^{\bar{x}} \beta(x) \cdot \left( \frac{f_a(x|a)}{f(x|a)} \right) f(x|a) dx \\ &= \int_0^{\hat{x}} -\beta(x) \left( \frac{f_a(x|a)}{f(x|a)} \right) f(x|a) dx + \int_{\hat{x}}^{\bar{x}} -\beta(x) \left( \frac{f_a(x|a)}{f(x|a)} \right) f(x|a) dx \\ &> \int_0^{\hat{x}} -\beta(x) \left( \frac{f_a(\hat{x}|a)}{f(\hat{x}|a)} \right) f(x|a) dx + \int_{\hat{x}}^{\bar{x}} -\beta(x) \left( \frac{f_a(\hat{x}|a)}{f(\hat{x}|a)} \right) f(x|a) dx \\ &= \left( \frac{f_a(\hat{x}|a)}{f(\hat{x}|a)} \right) \int_0^{\bar{x}} -\beta(x) f(x|a) dx = 0 \end{aligned} \quad (69)$$

Now, given a risky debt contract  $\delta(x) = \min\{x, D\}$ , the optimal effort  $a_n^*$  satisfies the first-order condition:  $\frac{T(D, a_n^*)}{\partial a} = 0$ . After the effort is sunk, at the renegotiation stage the entrepreneur's offer is an efficient ambiguity-sharing contract, which satisfies 35 given  $a_n^*$ . Finally,  $D^*$  satisfies the IR condition.  $\square$

## References

- Bewley, Truman F., 1989, Market innovation and entrepreneurship: A knightian view, Cowles Foundation Discussion Paper 905.
- Cornelli, Francesca, and Oved Yosha, 2003, Stage financing and the role of convertible securities, *Review of Economic Studies* 70, 1–32.
- Cover, Thomas M., and Joy A. Thomas, 2006, *Elements of Information Theory* (Wiley Science).
- Dewatripont, Mathias, Patrick Legros, and Steven A. Matthews, 2003, Moral hazard and capital structure dynamics, *Journal of the European Economic Association* 1, 890–930.
- Dicks, David, and Paolo Fulghieri, 2015, Ambiguity, disagreement, and allocation of control in firms, Working paper, SSRN.
- Dicks, David, and Paolo Fulghieri, 2016, Innovation waves, investor sentiment, and mergers, Working paper, SSRN.
- Dupuis, Paul, and Richard S. Ellis, 1997, *A Weak Convergence Approach to the Theory of Large Deviations* (John Wiley & Sons, Inc.).
- Ellsberg, Daniel, 1961, Risk, ambiguity, and the Savage axioms, *The Quarterly Journal of Economics* 75, 643–669.
- Fudenberg, Drew, and Jean Tirole, 1990, Moral hazard and renegotiation in agency contracts, *Econometrica* 58, 1279–1319.
- Garlappi, Lorenzo, Ron Giammarino, and Ali Lazrak, 2017, Ambiguity and the corporation: Group decisions, time inconsistency and underinvestment, *Journal of Financial Economics* 125, 417–433.
- Gilboa, Itzhak, and David Schmeidler, 1989, Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics* 18, 141–153.
- Hansen, Lars P., and Thomas J. Sargent, 2001, Robust control and model uncertainty, *American Economic Review* 91, 60–66.

- Hellmann, Thomas, 2006, IPOs, acquisitions, and the use of convertible securities in venture capital, *Journal of Financial Economics* 81, 649–679.
- Hermalin, Benjamin E., and Michael L. Katz, 1991, Moral hazard and verifiability: The effects of renegotiation in agency, *Econometrica* 59, 1735–1753.
- Innes, Robert D., 1990, Limited liability and incentive contracting with ex-ante action choices, *Journal of Economic Theory* 52, 45–67.
- Izhakian, Yehuda, and Jaime Zender, 2017, Ambiguity and disclosure in the principal-agent model, Working paper, SSRN.
- 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, 281–315.
- Kellner, Christian, 2017, The principal-agent problem with smooth ambiguity, *Review of Economic Design* 21, 83–119.
- Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2005, A smooth model of decision making under ambiguity, *Econometrica* 73, 1849–1892.
- Knight, Frank H., 1921, *Risk, Uncertainty and Profit* (Dover Classics).
- Lopomo, Giuseppe, Luca Rigotti, and Chris Shannon, 2011, Knightian uncertainty and moral hazard, *Journal of Economic Theory* 146, 1148–1172.
- Luenberger, David, 1969, *Optimization by Vector Space Methods* (John Wiley & Sons).
- Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, 2006, Ambiguity aversion, robustness, and the variational representation of preferences, *Econometrica* 74, 1447–1498.
- Malenko, Andrey, and Anton Tsoy, 2018, Asymmetric information and security design under Knightian uncertainty, SSRN working paper.
- Matthews, Steven A., 2001, Renegotiating moral hazard contracts under limited liability and monotonicity, *Journal of Economic Theory* 97, 1–29.

- Miao, Jianjun, and Alejandro Rivera, 2016, Robust contracts in continuous time, *Econometrica* 84, 1405–1440.
- Ortner, Juan, and Martin C. Schmalz, 2018, 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, 1257–1265.
- Schmidt, Klaus M., 2003, Convertible securities and venture capital finance, *Journal of Finance* 58, 1139–1166.
- Strzalecki, Tomasz, 2011, Axiomatic foundations of multiplier preferences, *Econometrica* 79, 47–73.
- Szydlowski, Martin, 2012, Ambiguity in dynamic contracts, Working paper, SSRN.
- Uppal, Raman, and Tan Wang, 2003, Model misspecification and underdiversification, *Journal of Finance* 58, 2465–2486.
- Wilson, Robert, 1968, The theory of syndicates, *Econometrica* 36, 119–132.