Risk Sharing, Risk Shifting and the Role of Convertible Debt

Saltuk Ozerturk∗
Department of Economics, Southern Methodist University

Abstract
This paper considers a financial contracting problem between a risk neutral entrepreneur and a risk averse investor. Once the venture is started, the entrepreneur chooses an action that determines the riskiness of the venture’s payoff. When action choice is contractible, the optimal risk sharing consideration under limited liability calls for a pure debt contract and the low risk action is adopted. When the action choice is not contractible, due to the risk shifting problem implementing the low risk action requires a deviation from the optimal risk sharing. I focus on situations where despite this deviation, the risk averse investor prefers to implement the low risk action and show that a convertible debt contract outperforms pure debt, pure equity and any mixture of debt and equity.

JEL CLASSIFICATION NUMBERS: D23, G24, G32

KEYWORDS: Convertible Debt, Second Order Stochastic Dominance, Financial Contracting.

REVISED & RESUBMITTED to JOURNAL OF MATHEMATICAL ECONOMICS

∗Assistant Professor of Economics, Southern Methodist University, Department of Economics, 3300 Dyer Street, Suite: 301, 75275 Dallas, TX. Tel: 214-768 33 44, Fax: 214-768 18 21, e-mail: ozerturk@mail.smu.edu. I am grateful to an anonymous referee for extremely useful comments. I would also like to thank Charles A.Wilson, Alberto Bisin, Bogachan Celen, Boyan Jovanovic, Douglas Gale, Kyle Hyndman, Levent Kockesen, Santanu Roy and seminar participants at New York University and Southern Methodist University for their helpful comments. The usual disclaimer applies.
1 Introduction

Financing a young entrepreneurial firm with a risky business plan is subject to important informational and incentive problems. Rather than more common instruments like debt or equity, investors who provide financing for entrepreneurial firms typically hold a convertible debt claim. In a recent empirical study on financial contracting by Kaplan and Stromberg (2003), convertible securities account for over 90% of all financing agreements in their sample of start-up firms. Previous work by Sahlman (1990) and Gompers (1997) also report the extensive use of convertible debt in venture capital backed entrepreneurial firms.

A convertible debt contract combines the properties of debt and equity. The conversion option gives the claimholder the right to convert the debt claim into company’s equity. This paper offers an explanation on why convertible debt can be superior to pure debt, pure equity and mixed debt-equity in the venture capital context. I describe a financial contracting problem where a risk neutral entrepreneur finances his venture by funds provided by a risk averse financier/venture capitalist. The admissible sharing rules on the final payoff of the venture are constrained by a limited liability condition. Upon receiving the funds, the entrepreneur adopts a business strategy, which cannot be specified ex ante by the contract. This action choice determines the riskiness of the venture’s payoff. In particular, I consider a simple model with two possible actions where the high risk action \( a_H \) yields a payoff distribution which is a mean preserving spread of the distribution induced by the low risk action \( a_L \).

When the action choice is enforceable, under limited liability a pure debt contract achieves optimal risk sharing between parties. Furthermore, under enforceability of actions the financier prefers to enforce the low risk action. When the action choice is not enforceable, a debt contract induces the entrepreneur a preference for the high risk action due to its convex residual payoff. Therefore, implementing the low risk action requires a deviation from the optimal risk sharing arrangement provided by the debt contract. I focus on situations where despite this deviation from optimal risk sharing, implementing the low risk action makes the risk averse financier better off compared to opting for the high risk action. In this setting, a convertible debt contract outperforms pure debt, pure equity and any mixture of debt and equity. This result follows because a convertible debt contract combines two desirable properties in this problem. Its debt component assigns the whole payoff to the risk averse financier at the low end of payoff realizations and provides better insurance against the downside risk. The conversion into equity option, on the other hand, provides a convex payoff schedule for the financier at the upper end of payoff realizations, and corrects the entrepreneur’s high risk incentives arising from the debt portion of the
contract. This role can not be achieved by simple mixtures of debt and equity, since mixed debt-equity contracts also yield a concave payoff for the financier (and hence a convex one for the entrepreneur) and implement the high risk action.

My primary focus is to offer an explanation on why convertible debt can be superior to more traditional financial contracts such as pure debt, pure equity and mixed debt-equity. As mentioned, this superiority stems from the fact that (i) unlike an equity contract convertible debt gives the whole payoff to the risk averse financier below a certain payoff realization and improves risk sharing, (ii) unlike debt and mixed debt-equity, convertible debt assigns the financier a convex increasing payoff schedule at the upside and prevents the entrepreneur from increasing risk ex-post. For further insight, in an extension I analyze whether an optimal solution can exhibit these two properties. I show that these two properties can emerge provided that (i) the admissible contracts exhibit a monotonicity property, i.e., the financier’s payoff is constrained to be non-decreasing in the venture’s payoff (as in Innes (1990)), and (ii) the likelihood ratio $f_H/f_L$ of the density functions for the high and low risk distributions is first decreasing and then increasing. This regularity condition on the likelihood ratio $f_H/f_L$ is consistent with the mean preserving spread assumption, and is satisfied, for example, in the case of a normal distribution.

Since the risk sharing consideration between the risk averse financier and the risk neutral entrepreneur is a crucial aspect of the analysis, the assumption on the risk attitudes of the two parties deserves further comment. The financial contracting problem posed in this paper can be best placed in the venture capital industry context. A typical venture capitalist periodically raises what is called a venture capital fund from cash rich institutions such as pension funds and insurance companies. A venture capital fund has a lifetime of five to seven years, at the end of which the returns are distributed to the fund contributors. In that sense, the venture capitalist acts as a fund manager by choosing which ventures to invest. The ability to raise new capital for future funds depends on the performance of the earlier investments. Sahlman (1990) notes that if a venture capitalist’s fund suffers huge losses or even in cases of moderate failures, the chances for raising new capital for the next fund are very limited. As pointed out by Chemmanur and Fulghieri (1999), in a given fund cycle the venture capitalist can only invest in a few ventures, all of which are quite risky, and therefore the capitalist’s investment portfolio remains poorly diversified. They also consider a risk averse venture capitalist and argue that the compensation scheme of a venture capitalist from managing a fund involves significant penalties for failures, thereby inducing risk averse behavior. Furthermore, the success or failure of a

---

1 Sahlman (1990) analyzes a sample of 383 venture capital investments and reports that 35% of all projects yielded a net loss and another 50% were only moderately successful.

2 For a similar discussion for the risk averse behavior of banks, see Diamond (1984).
particular project can significantly affect the reputation of the venture capitalist who made the decision to invest in that venture, again leading to risk aversion.3

As in Innes (1990) and Chemmanur and Fulghieri (1999), I consider a risk neutral entrepreneur with the following justification. Unlike the ‘agent’ of a standard principal-agent model (typically interpreted as a risk averse employee of the principal), an entrepreneur who quits a well paying job to pursue a fortune by launching a new company does not seem to be exhibiting risk aversive behaviour. Furthermore, what an entrepreneur loses from a failing venture is considerably less than what a venture capitalist loses. For a typical entrepreneur, even receiving funds for the venture can work as a badge of success, and having been in charge of a company, even if it eventually fails, is a valuable experience for a future career, especially in a growing young industry.4

Related Literature. Previous explanations of convertible securities have focused either on the efficient allocation of control rights paradigm or on an effort type moral hazard problem. Berglof (1994) provides a model where control refers to the right to bargain with an outside party bidding for the venture and shows that convertible security allocates the control to the party who maximizes the joint surplus of the entrepreneur and the financier. Another control based explanation is Marx (1998) where a mixture of debt and equity dominates pure debt and pure equity in giving the financier the efficient liquidation incentives. However, as Gompers (1997) convincingly argues, allocation of cash flow rights and allocation of control rights can be separated by use of covenants and explicit contractual clauses. Indeed, Gompers (1997) documents the frequent use of covenants that give investors control rights. We take the view that such control rights are somewhat independent from the design of financial instrument and the primary purpose of convertible security is more likely to be risk sharing and agency considerations, which is the focus of this paper. Cornelli and Yosha (2003) show that conversion into equity option can be desirable, because it may prevent the entrepreneur from window dressing (short-termism) which does not contribute to the long term success of the venture. Trester (1998) shows that a financier’s conversion into equity option may prevent the entrepreneur from defaulting strategically and walking away from the venture. However, it seems that what

---

3 Consistent with this view, Sahlman (1990) notes that when valuing a company, a venture capitalist computes the present value of a company by applying a very high discount rate, usually in the range of 40% to 60%.

4 The following remarks of Joseph Park, the founder of Kozmo.com, illustrates this point: ‘Let’s say I completely failed in 6 months after launching the company and lost all the money I raised. So what? I will have an impressive resume to apply to a business school’ (from the documentary film e-Dreams). Kozmo.com was a venture capital driven online company that promised free one-hour delivery of anything from DVD rentals to Starbucks coffee. After raising about $280 million, the company had to shut down its operations in 2001. Perhaps ironically, it is now a widely studied example of the dot-com excess and made Joseph Park a celebrity in business school case studies.
prevents an entrepreneur from walking away from the venture is the fact that her stakes are vested over time and only become liquid after a certain period of time. Furthermore, many agreements give the investor the right to purchase a departing entrepreneur’s share at a low price (see Sahlman 1990). Schmidt (2003) shows that convertible debt gives efficient investment incentives when the entrepreneur and investor move sequentially in a double moral hazard type problem.

Another related paper is Innes (1990) who considers an effort type agency model, where higher levels of costly effort by the entrepreneur induces better payoff distributions in the sense of the monotone likelihood ratio property. Innes first allows for non-monotonic sharing rules and shows that the optimal contract punishes the entrepreneur by giving the investor all the realized payoff below a threshold, whereas the entrepreneur is awarded by receiving all the payoff over this threshold. To rule out this non-monotonic optimal contract, Innes then imposes a monotonicity constraint on the admissible sharing rules and establishes that under this additional constraint a simple debt contract emerges as optimal. Different than Innes who assumes risk neutrality for both parties, the risk sharing consideration plays an important role in my model. Furthermore, the entrepreneur’s action choice determines the riskiness of the venture’s payoff, but does not affect its expected value.

The plan of the paper is as follows. Section 2 presents the model. Section 3 provides the benchmark case when the action choice of the entrepreneur is enforceable. Section 4 considers the case when action choice is not observable. Section 5 concludes.

2 The Model

There are three dates, $t = 0, 1, 2$. There is an entrepreneur (henceforth EN) who owns a venture idea. The venture requires a fixed investment of $K$ at date 0. The EN has no wealth of his own and relies on a financier/investor (henceforth FI) to provide the investment capital. This financier can be thought as a venture capitalist focusing on young entrepreneurial firms. At date 2, the venture generates a random payoff $\tilde{y}$. The realizations of the random variable $\tilde{y}$ (that I denote with $y$) are drawn from a support $[0, \infty)$. The EN is risk neutral and maximizes expected wealth. The FI maximizes a strictly concave VNM utility function $v(\cdot)$ with $\lim_{y \downarrow 0} v'(y) = \infty$.

The distribution of the venture’s payoff $\tilde{y}$ depends on the action that EN chooses at date 1. This action can be thought as the business strategy employed by EN upon receiving the required funds. For simplicity, I consider two mutually exclusive actions denoted by $a_L$ and $a_H$. Formally, for $i \in \{H, L\}$, let $F_i(y)$ denote the distribution function from action $a_i$ with a continuously differentiable density $f_i$. Following Rothschild and Stiglitz (1970), the following assumption ranks the riskiness of the two distributions under two actions.
**Assumption 1.** The payoff distribution $F_H$ is a mean preserving spread of the payoff distribution $F_L$, i.e.,

$$
\int_0^x (F_H - F_L) \, dy \geq 0 \text{ for all } x > 0, \text{ and }
$$

$$
\int_0^\infty y f_L dy = \int_0^\infty y f_H dy = \mu. \quad (1b)
$$

The above assumption says that the action $a_H$ yields a riskier payoff distribution than action $a_L$. In the context of a start-up company operating in an innovative industry, the riskiness of the business plan is an important determinant of failure or success. Among the many possible ways to increase risk in start-up environments, the most common ones are: rushing the product to the market although further testing is warranted, changing the scope of venture’s operations and drifting into uncharted territory, insisting on a very ambitious design feature and thus increasing technical risk.\(^5\)

While the realization of the venture’s payoff is observable and contractible, I assume that the action $a_i$ is not contractible, i.e., the two parties cannot write an enforceable contract clause at date 0 that dictates EN to choose a particular action. A sharing rule (a financial contract) is an integrable function $s(y) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ which specifies the payment to FI for each payoff outcome $y$. If the realized payoff is $y$, then the FI receives $s(y)$ and EN as the residual claimant receives $y - s(y)$. As in Innes (1990), I assume that the sharing rule $s(y)$ must exhibit limited liability so that $0 \leq s(y) \leq y$ for all $y$. This limited liability constraint implies that EN cannot be forced to pay FI more than what the venture generates ($s(y) \leq y$) and FI cannot be forced to make an additional transfer once the venture’s payoff is realized ($s(y) \geq 0$).

I consider a setting with many entrepreneurs seeking funds for their ventures, but only a few financiers/venture capitalists who can provide financing for young entrepreneurial firms. The venture capitalists typically specialize in certain industries such as biotechnology and telecommunications and their industry specific expertise also serves as an entry barrier for less specialized financiers (Gompers 1997). Furthermore, especially at the initial stage of a venture, the business plan and the skills of an entrepreneur are completely untested, giving EN not much bargaining power. Accordingly, I assume that FI has all the bargaining power and chooses the sharing rule that gives her the maximal expected utility subject to a participation constraint for EN, any required incentive compatibility constraint and the limited liability constraint. The assumption that FI extracts all the surplus does not affect the qualitative results.\(^6\)

---

\(^5\)I do not consider the possibility of learning about the venture’s prospects over time. For a model with this feature, see Bergemann and Hege (1998).

\(^6\)The formulation that maximizes FI’s expected utility subject to giving EN his reservation utility is merely a tool to describe the Pareto optimal solution for any given reservation utility level for EN.
3 Action choice enforceable

As a benchmark, this section considers the case when the action choice \( a_i \) is enforceable. The optimal contract in this case is determined only by risk sharing considerations. To state the optimal contracting problem formally, define

\[
U_i(s(.)) \equiv \int_0^\infty (y - s(y)) f_i dy \text{ for } i \in \{H, L\},
\]

as the expected payoff for EN from adopting action \( a_i \) under a sharing rule \( s(y) \). An optimal contract that prescribes EN to undertake action \( a_i \) is then the solution to a programming problem in which the sharing rule \( s(y) \) is chosen to maximize the expected utility of FI,

\[
V_i(s(.)) \equiv \int_0^\infty v(s(y)) f_i dy
\]

subject to a participation constraint for EN,

\[
U_i(s(.)) \geq w > 0,
\]

and the limited liability constraint

\[
0 \leq s(y) \leq y \text{ for all } y.
\]

I refer to the above problem as (P1). The following proposition establishes that the optimal risk sharing consideration under limited liability calls for a pure debt contract.

**Proposition 1a.** If action choice \( a_i \) is enforceable, the optimal risk sharing contract that prescribes EN to undertake action \( a_i \) is a pure debt contract

\[
s_i(y) = \text{Min}(y, m_i),
\]

where the face value \( m_i \) of debt is uniquely determined by EN’s participation constraint and solves

\[
m_i - \int_0^{m_i} F_i dy = \mu - w.
\]

**Proof:** In an optimal solution to (P1), EN’s participation constraint (4) holds as an equality. Under a pure debt contract \( s_i(y) = \text{Min}(y, m_i) \) that prescribes \( a_i \), the participation constraint becomes

\[
U_i = \int_{m_i}^\infty (y - m_i) f_i dy = \mu - m_i + \int_0^{m_i} F_i dy = w.
\]

and yields (6b). I now show that \( s_i(y) = \text{Min}(y, m_i) \) is the optimal solution to (P1). Suppose, contrary to our claim, that there is an alternative sharing rule \( \hat{s}_i(y) \), different than the debt contract \( s_i(y) \) in (6a-6b), which is a solution to (P1). This would imply

\[
\int_0^\infty v(\hat{s}_i(y)) f_i dy \geq \int_0^\infty v(s_i(y)) f_i dy.
\]
Let \( \hat{G} \) and \( \hat{F}_i \) denote the distribution functions of \( \hat{s}_i(y) \) and \( s_i(y) \), respectively. Since \( v(.) \) is a strictly concave increasing function, the inequality in (7a) implies that \( \hat{G} \) second order stochastically dominates \( \hat{F}_i \) (see Hadar and Russell (1969)), and therefore we have

\[
\int_0^x \left( \hat{F}_i - \hat{G} \right) \, dy \geq 0 \quad \text{for all } x > 0. \tag{7b}
\]

Note that the distribution function \( \hat{F}_i(.) \) of \( s_i(.) \) is given by

\[
\hat{F}_i(y) = F_i(y) \text{ for } y < m_i \text{ and } \hat{F}_i(y) = 1 \text{ for } y \geq m_i.
\]

To obtain a contradiction, suppose that for realizations \( y \in [0, m_i] \) our alternative optimal sharing rule \( \hat{s}_i(y) \) is different than the debt contract \( s_i(y) \) in (6a-6b). With the limited liability restriction in (5), this implies that \( \hat{s}_i(y) \leq y \) for all \( y \in [0, m_i] \) and \( \hat{s}(y) < y \) for some \( y \in [0, m_i] \). But then we have

\[
\hat{G}(y) \quad \geq \quad \hat{F}_i(y) = F_i(y) \text{ for all } y \in [0, m_i] \text{ and }
\]

\[
\hat{G}(y) \quad > \quad \hat{F}_i(y) = F_i(y) \text{ for some } y \in [0, m_i],
\]

which further implies that

\[
\int_0^{m_i} \left( \hat{F}_i - \hat{G} \right) \, dy < 0, \tag{7c}
\]

and contradicts (7b). Accordingly, the proposed alternative optimal sharing rule \( \hat{s}_i(y) \) cannot be different than the debt contract \( s_i(y) \) in (6a-6b) for \( y \in [0, m_i] \), as otherwise it would be second order stochastically dominated by \( s_i(y) \). This argument establishes that an optimal solution to (P1) must have \( s(y) = y \) for \( y \in [0, m_i] \).

To prove that \( \hat{s}_i(y) \) cannot be different than the debt contract \( s_i(y) \) in (6a-6b) for \( y > m_i \) either, note that both \( \hat{s}_i(y) \) and \( s_i(y) \) must give FI an expected payoff of \( \mu - w \) and therefore we have

\[
E[\hat{s}_i(y)] = E[s_i(y)] \Rightarrow \int_0^\infty \left( \hat{F}_i - \hat{G} \right) \, dy = 0. \tag{8a}
\]

It is already shown that \( \hat{s}(y) = s_i(y) = y \) for \( y \in [0, m_i] \), and hence the equality in (8a) becomes

\[
\int_{m_i}^\infty \left( \hat{F}_i - \hat{G} \right) \, dy = 0. \tag{8b}
\]

Since \( \hat{F}_i(y) = 1 \) for \( y \geq m_i \), the equality in (8b) implies that \( \hat{G}(y) = 1 \) for \( y \geq m_i \). This proves that \( \hat{s}_i(y) = m_i \) for \( y \geq m_i \) as well. Accordingly, the unique solution to (P1) is given by \( s_i(y) \) in (6a-6b). \( Q.E.D. \)
One can write the expected utility of FI from holding a debt claim \( s_i(y) = \text{Min}(y, m_i) \) as

\[
V_i(s_i(y)) = \int_0^{m_i} v(y)f_i dy + v(m_i)(1 - F_i(m_i)) = v(m_i) - \int_0^{m_i} v'(y)F_i dy,
\]

where the second equality follows from integration by parts. In the above expression, the second term \( \int_0^{m_i} v'(y)F_i dy \) accounts for the reduction in FI's expected utility due to riskiness of the debt claim. For a given \( m_i \), this term is increasing in the riskiness of the distribution \( F_i \) and the risk aversion of FI. This is because under the high risk action \( a_H \) default is more likely and a more risk averse FI suffers more at the low end of payoff realizations. As a result, to compensate for this reduction one would expect the face value to be higher when the high risk choice \( a_H \) is adopted. I now prove that this is indeed the case and we have \( m_H > m_L \). Recall that the face value \( m_i \) is determined by the action \( a_i \) undertaken and EN's reservation wage \( w \). From (6b), we have

\[
m_L - m_H = \int_0^{m_L} F_L dy - \int_0^{m_H} F_H dy.
\]

For a contradiction, suppose that \( m_L \geq m_H \) and rewrite (10a) as

\[
m_L - m_H = \int_0^{m_H} (F_L - F_H) dy + \int_{m_H}^{m_L} F_L dy.
\]

By Assumption 1, the first term \( \int_0^{m_H} (F_L - F_H) dy \) \( \leq 0 \). Therefore, if \( m_L \geq m_H \) from (10b) we would have

\[
m_L - m_H \leq \int_{m_H}^{m_L} F_L dy.
\]

But this last inequality in (10c) is a contradiction, since \( F_L(y) < 1 \) for \( y \in [m_H, m_L] \) and hence \( \int_{m_H}^{m_L} F_L dy < m_L - m_H \). This argument establishes that \( m_L < m_H \). Using this observation, one can show that when the action choice is enforceable, FI strictly prefers the low risk action \( a_L \).

**Proposition 1b.** If the action choice is enforceable, FI strictly prefers the low risk action \( a_L \) to be undertaken by offering \( s_L(y) = \text{Min}(y, m_L) \).

**Proof.** Given \( m_L < m_H \), I show that \( s_L \) second order stochastically dominates \( s_H \). The desired result then follows from the strict concavity of \( v(.) \). Since \( m_L < m_H \), one can write equation (10a) as,

\[
\int_0^{m_L} (F_H - F_L) dy = \int_{m_L}^{m_H} (1 - F_H) dy.
\]
Let $\hat{F}_i$ denote the distribution function for $s_i$. We have $\hat{F}_i(y) = F_i(y)$ for $y < m_i$ and $\hat{F}_i(y) = 1$ for $y \geq m_i$. Since the expected payoff is the same under both $s_L(y)$ and $s_H(y)$, equation (11) implies

$$\int_{0}^{x} \left( \hat{F}_L - \hat{F}_H \right) dy < 0 \text{ for } x < m_H \text{ and}$$

$$\int_{0}^{x} \left( \hat{F}_L - \hat{F}_H \right) dy = 0 \text{ for } x \geq m_H,$$

which establishes that $s_L$ second order stochastically dominates $s_H$. Q.E.D.

The analysis in this section established that when the action choice is enforceable, for a given action $a_i$ (and hence a given distribution $F_i$) a pure debt contract provides the best possible insurance (as much as limited liability allows) for the risk averse FI by allocating her all the realized payoff at the downside. Furthermore, under enforceability FI strictly prefers to enforce the low risk action $a_L$. After characterizing this benchmark case under enforceability, I now analyze the case when action choice is not enforceable.

4 Action choice not enforceable

4.1 Risk Shifting

Suppose now the action choice of EN is not enforceable, and EN chooses the action that maximizes his expected payoff given the sharing rule specified at date 0. A well known result in the financial contracting literature is that under a risky debt contract the risk neutral EN has a preference for high risk (see, for example, Jensen and Meckling (1976), Green (1984)). To illustrate the risk shifting incentives in our framework, consider any pure debt contract with a repayment obligation $m$. From (6c), EN’s expected payoff from choosing action $a_i$ under such a debt contract is given by

$$U_i = \mu - m + \int_{0}^{m} F_i dy \text{ for } i \in \{H, L\}. \tag{12}$$

The change in EN’s payoff from switching to the high risk action $a_H$ under the debt contract can then be written as

$$U_H - U_L = \int_{0}^{m} (F_H - F_L) dy \geq 0 \text{ for all } m > 0. \tag{13a}$$

Therefore, a pure debt contract, which provides optimal risk sharing under enforceability, implements the high risk action $a_H$ when action choice is not enforceable. The reason behind the entrepreneur’s preference for high risk is the convexity of the residual payoff for EN under a pure debt contract. To see this more transparently,
consider any concave sharing rule \( s^c(y) \) (including the pure debt contract) that yields a convex residual payoff \( y - s^c(y) \) for EN. By Assumption 1, we have

\[
U_H - U_L = \int_0^\infty (y - s^c(y)) (f_H - f_L) \, dy = \int_0^\infty s^c(y) (f_L - f_H) \, dy \geq 0. \tag{13b}
\]

and hence EN’s preference for high risk extends to any concave sharing rule \( s^c(y) \). Accordingly, a mixed debt-equity contract of the form \( s(y) \equiv \text{Min}(y, m + \pi(y - m)) \) with \( m > 0 \) and \( \pi \geq 0 \), which gives the FI a share of the upside of the venture, is also concave, and implements the high risk action \( a_H \). In a mixed-debt equity contract, EN’s payoff from switching to \( a_H \) is given by

\[
U_H - U_L = (1 - \pi) \int_0^m (F_H - F_L) \, dy \geq 0. \tag{13c}
\]

### 4.2 Implementing the Low Risk Action

Due to the risk shifting problem, implementing \( a_L \) requires FI to deviate from optimal risk sharing achieved by the debt contract and satisfy an additional incentive compatibility constraint. Formally, the problem that FI needs to solve to implement \( a_L \) optimally can be stated as follows.

\[
\text{Max}_{s(\cdot)} \quad V_L(s(y)) = \int_0^\infty v(s(y)) f_L \, dy \tag{P2}
\]

\[
\text{s.t.} \quad \int_0^\infty (y - s(y)) f_L \, dy \geq w > 0, \tag{14a}
\]

\[
\int_0^\infty (y - s(y)) (f_L - f_H) \, dy = \int_0^\infty s(y) (f_H - f_L) \, dy \geq 0, \tag{14b}
\]

and the limited liability constraint \( 0 \leq s(y) \leq y \) for all \( y \). I refer to the above problem as (P2). Before proceeding, a discussion is in order. Note again that implementing \( a_L \) by satisfying (14b) involves an agency cost relative to the utility level achieved by \( s_L(y) = \text{Min}(y, m_L) \), since it requires a deviation from the optimal risk sharing rule. On the other hand, FI can always opt for offering the debt contract \( s_H(y) = \text{Min}(y, m_H) \) to implement \( a_H \) and ensure an expected utility

\[
V_H(s_H(y)) = v(m_H) - \int_0^{m_H} v'(y) F_H \, dy.
\]

One question is whether FI is better off from satisfying (14b) and solving (P2) compared to implementing \( a_H \) and receiving \( V_H(s_H(y)) \). While the answer depends on the relative riskiness of \( F_H \) and \( F_L \), the only restriction I have is a mean preserving spread condition, which is too weak to mathematically characterize a simple condition that ensures FI is better off from implementing \( a_L \). I focus on situations where
implementing the low risk action brings a higher expected utility to the risk averse FI than the maximum she can get under the high risk action. This would be the case if \( F_H \) is risky enough and/or FI is risk averse enough so that \( V_H(s_H(y)) \) is low enough. As is common in the agency literature, in what follows I consider the case that despite the deviation from optimal risk sharing and associated agency cost, the second best involves implementing \( a_L \) rather than opting for \( a_H \).\(^7\)

It should be noted that a sharing rule that satisfies (14b) and implements \( a_L \) can yield the risk averse FI a higher expected utility than \( V_H(s_H(y)) \), even if it is not an optimal solution to (P2). In particular, FI may be better off from implementing \( a_L \) even with a pure equity contract \( s_E(y) = \pi y \) which satisfies (14b), but is not necessarily an optimal solution to (P2). The following example, developed and analyzed more completely in the Appendix, illustrates this point. Let \( v(y) = 2y^{1/2} \) and specify the two distributions as \( F_L(y) = \frac{y^2}{2} \) for \( y \in [0, 2] \) and \( F_H(y) = \frac{1}{4} + \frac{y}{4} \) for \( y \in [0, 2) \) and \( f_H(2) = \frac{1}{4} \). Set EN’s reservation wage to \( w = \frac{1}{4} \). One can compute that (see the Appendix for a general derivation), the optimal debt contract that implements \( a_H \) is given by \( s_H(y) = \text{Min}(y,m_H = 1.26) \) which yields FI an expected utility \( V_H(s_H(y)) = 1.45 \). On the other hand, an equity contract \( s_E(y) = 0.75y \) gives EN a reservation wage \( w = \frac{1}{4} \), implements \( a_L \) and yields FI an expected utility \( V_L(s_E(y)) = 1.64 \).

### 4.3 Superiority of Convertible Debt over Pure Debt and Equity

My main objective is to provide an explanation on why convertible debt can be superior to more commonly observed financial contracts such as pure debt, pure equity and mixed-debt equity. The ‘convertible debt’ contract can be described as a sharing rule

\[
s(y; m, \pi) \equiv \max \{ \min \{ y, m \}, \pi y \} \quad \text{for} \ m \geq 0 \ \text{and} \ \pi \in [0, 1)
\]  

(15)

The above contract specifies a payoff realization \( m \) below which FI receives all the realized payoff. Therefore, \( m \) again serves as the face value of FI’s debt claim. The conversion into equity option is described by the share \( \pi \) of venture’s equity. Upon realization of the venture’s payoff, FI has the option to exchange the debt claim \( m \) for a share \( \pi \) of the venture’s equity. This conversion into equity option is exercised for payoff realizations \( y \geq m/\pi \), whereas for \( y < m/\pi \) FI retains the debt claim. It should be noted that the pure equity and the pure debt contracts are special cases in this

\(^7\)I am grateful to an anonymous referee for raising this question. If we have \( V_L(s^*(y)) < V_H(s_H(y)) \) at an optimal solution \( s^*(y) \) to (P2), then the second best contract is \( s_H(y) \) and the high risk action \( a_H \) should be implemented. This would be similar to an agency model with costly hidden effort where the agency cost of inducing the agent the high effort level is higher than the efficiency benefits on expected output and as a result the principal has to settle for the low effort.
family. The contract $s(y; m = 0, \pi)$ corresponds to a pure equity contract, whereas $s(y; m, \pi = 0)$ corresponds to a pure debt contract. The figure below illustrates FI’s payoff schedule from a convertible debt contract $s(y; m, \pi)$.

![Figure 1. The convertible debt contract](image)

A convertible debt contract has two desirable properties compared to pure debt and pure equity in the context of (P2). Unlike a pure equity contract, the debt component of convertible debt assigns the whole payoff to the risk averse FI at the low end of payoff realizations and provides better insurance at the downside. Furthermore, unlike a pure debt contract (or mixed debt-equity), the conversion into equity option of convertible debt creates a convex payoff schedule for FI at the upper end of payoff realizations and corrects EN’s high risk incentives arising from FI’s debt claim. Accordingly, the convertible debt contract provides better insurance by its debt component, while eliminating EN’s preference for high risk with its conversion into equity component.

I now formalize the above argument and show that convertible debt outperforms pure equity and pure debt in problem (P2). Let me rewrite (P2) by restricting attention to the class of sharing rules $s(y; m, \pi)$ for $m \geq 0$ and $\pi \in [0, 1)$. The FI’s expected utility from $s(y; m, \pi)$ is given by

$$V_L(s(y; m, \pi)) = \int_0^m v(y)f_Ldy + \int_0^\pi v(m)f_Ldy + \int_{m}^{\infty} v(\pi y)f_Ldy.$$  \hspace{1cm} (16)

The participation constraint for EN takes the form

$$U_L(s(y; m, \pi)) = \int_0^m \pi (y - m) f_Ldy + \int_{m}^{\infty} (1 - \pi) y f_Ldy \geq w.$$  \hspace{1cm} (17)

Finally, define $W(s(y; m, \pi)) \equiv U_L(s(y; m, \pi)) - U_H(s(y; m, \pi))$. Then the incentive compatibility constraint in (14b) becomes

$$W(s(y; m, \pi)) \equiv \int_0^m (y - m)(f_L - f_H)dy + (1 - \pi) \int_{m}^{\infty} y(f_L - f_H)dy \geq 0.$$  \hspace{1cm} (18)
Within the family of sharing rules \( s(y; m, \pi) \), FI chooses \( m \geq 0 \) and \( \pi \in [0,1) \) to maximize \( V_L(s(y; m, \pi)) \) subject to (17) and (18). The proposition below formalizes that a non-degenerate convertible debt contract \( s(y; m > 0, \pi > 0) \) outperforms pure debt and pure equity in (P2).

**Proposition 2.** A convertible debt contract outperforms pure debt and pure equity in (P2).

**Proof.** The preceding analysis ruled out a pure debt contract, since it implements the high risk action and violates (18). I now show that \( s(y; m > 0, \pi > 0) \) outperforms the equity contract \( s(y; m = 0, \pi > 0) \). Define \( m^* \) by \( U(m^*, 0) = w \). Since \( U \) is strictly decreasing in \( m \) and \( \pi \), one can define a function \( \pi : [0, m^*] \to [0,1] \) by \( U(m, \pi(m)) \equiv w \). Formally, the function \( \pi(m) \) satisfies

\[
U(m, \pi(m)) \equiv \int_m^{m^*} (y - m) f_L dy + \int_{m^*}^{\infty} (1 - \pi(m)) y f_L dy = w
\]

and implicit differentiation yields

\[
\pi'(m) = -\frac{\int_m^{m^*} f_L dy}{\int_{m^*}^{\infty} y f_L dy} < 0,
\]

which implies that for a given reservation payoff \( w \) for EN, increasing FI's debt claim requires decreasing the equity claim, so that the participation constraint continues to hold as an equality.

Note that a pure equity contract also implements the low risk action and therefore we have \( W(0, \pi) = 0 \). To show that a solution with \( m > 0 \) is superior, one needs to show that for \( m > 0 \) sufficiently small, we have

\[
\frac{dV(m, \pi(m))}{dm} > 0 \quad \text{and} \quad \frac{dW(m, \pi(m))}{dm} \geq 0.
\]

The idea is to start with \( m = 0 \) which satisfies (18) and then to find a sufficiently small \( m > 0 \) that makes FI better off yet still implementing the low risk action. The equity share \( \pi(.) \) defined by the binding participation constraint above adjusts accordingly. Note that

\[
\frac{dV(.)}{dm} = \int_m^{m^*} f_L dy \left( 1 - \frac{\int_{m^*}^{\infty} v'(\pi y) y f_L dy}{\int_{m^*}^{\infty} y f_L dy} \right) > \int_m^{m^*} f_L dy \left( 1 - \frac{\int_{m^*}^{\infty} y f_L dy}{\int_{m^*}^{\infty} y f_L dy} \right) = 0
\]

where the inequality follows from the fact that \( v'(m) \) converges to a large number for \( m \) sufficiently small and therefore the ratio \( v'(\pi y)/v'(m) \) approaches to zero.
Similarly, straightforward algebra yields

\[
\frac{dW(.)}{dm} = \int_m^\infty f_H dy - \int_m^\infty f_L dy \left[ \int_m^\infty y f_H dy - \int_m^\infty y f_L dy \right] > \int_m^\infty (f_H - f_L) dy > 0
\]

for \( m \) small enough. Note that we use the fact that for sufficiently small \( m \), Assumption 1 implies that \( f_H(y) > f_L(y) \) for all \( y \in (m, \frac{m}{\pi}) \). \( Q.E.D. \)

### 4.4 Characterizing an Optimal Solution to (P2)

The preceding analysis established that a convertible debt contract is superior to pure debt and pure equity in (P2) due to its two properties: (i) at the lower end of payoff realizations it assigns the whole payoff to the risk averse FI and improves risk sharing, (ii) at the upper end of payoff realizations, it creates convexity in FI’s payoff schedule and corrects EN’s high risk incentives. For further insight, in this section I provide an analysis of the optimal solution to (P2) to see whether these two properties can emerge in a general solution as well.

As I show shortly, the behaviour of an optimal solution to (P2) depends on the likelihood ratio of the density functions \( f_H/f_L \). The mean preserving spread assumption (Assumption 1) alone does not preclude the possibility of considerable local variation in the relative magnitudes of \( f_L \) and \( f_H \). Consequently, to obtain a reasonably well-behaved solution, one needs to impose additional regularity conditions on \( f_H/f_L \). I impose the following regularity condition on \( f_H/f_L \), which is consistent with Assumption 1.

**Assumption 2.** There exists a \( y^* > 0 \) such that \( f_H/f_L \) is strictly decreasing for \( y < y^* \) and strictly increasing for \( y > y^* \).

The above assumption on \( f_H/f_L \) holds, for example, in the case of two normal distributions with the same mean and different variances. Along with the limited liability restriction in (5), I follow Innes (1990) and also impose the following monotonicity restriction on the admissible sharing rules.

**Assumption 3.** FI’s payoff \( s(y) \) must be non-decreasing in the venture’s payoff, i.e., \( s(y + \varepsilon) \geq s(y) \) for all \( (y, \varepsilon) \in \mathbb{R}_+^2 \).

As in Innes (1990), the above assumption can be justified as follows (for a further discussion, see also Bolton and Dewatripont (2005, page 164)). Suppose that Assumption 3 is violated and there is a segment of payoff realizations such that \( s(y) \) is

---

8 Consider two normal distributions with the same mean \( \mu \) and variances \( \sigma_H^2 \) and \( \sigma_L^2 \) with \( \sigma_H > \sigma_L \). Then the likelihood ratio \( f_H/f_L \) is strictly decreasing for \( y < \mu \) and strictly increasing for \( y > \mu \).
strictly decreasing in $y$. In that segment, EN would strictly gain by borrowing money at par from another source and thus marginally boost the venture’s performance. If this kind of costless borrowing goes undetected, EN would have an incentive to engage in an arbitrage activity by borrowing at any decreasing segment of $s(y)$, artificially boosting the venture’s performance and reducing FI’s payoff.\footnote{As noted by Innes(1990), since EN cannot be forced to borrow, the limited liability constraint $s(y) \leq y$ still applies here.}

I now proceed with the characterization of an optimal solution to (P2) under these two additional assumptions. Let $\lambda \geq 0$ and $\delta \geq 0$ denote the Lagrange multipliers in (P2) associated with EN’s participation constraint (14a) and the incentive compatibility constraint (14b), respectively. To save notation, I do not introduce additional Lagrange multipliers for the limited liability constraint (5) and the monotonicity constraint (Assumption 3), but check separately whether they are binding. The Lagrangean for (P2) takes the form

$$
\mathcal{L} = \int_0^\infty v(s(y))f_L dy + \lambda \left( \int_0^\infty (y - s(y)) f_L dy - w \right) + \delta \left( \int_0^\infty (y - s(y))(f_L - f_H) dy \right).
$$

An optimal solution to (P2) must give EN an expected payoff no more than the reservation wage $w$ and therefore $\lambda > 0$. Also note that $\delta = 0$ is not possible, as otherwise pure debt would be optimal, which we know violates (14b). Therefore, both constraints are binding, and $\lambda > 0, \delta > 0$. Pointwise differentiation of $\mathcal{L}$ with respect to $s(y)$ and using the limited liability constraint as a boundary condition, one obtains the following first order condition:

$$
v'(s^*(y)) \leq \lambda + \delta \left[ 1 - \frac{f_H(y)}{f_L(y)} \right] \text{ if } s^*(y) < y,
$$

$$
v'(s^*(y)) \geq \lambda + \delta \left[ 1 - \frac{f_H(y)}{f_L(y)} \right] \text{ if } s^*(y) > 0.
$$

The binding participation constraint (14a) with $w > 0$ requires that $s^*(y) < y$ for some $y$. Therefore, one can define

$$
y_1 \equiv \inf \left\{ y > 0 : v'(s(y)) \leq \lambda + \delta \left[ 1 - \frac{f_H(y)}{f_L(y)} \right] \right\}, \tag{19}
$$

$$
y_2 \equiv \sup \left\{ y > 0 : v'(s(y)) \leq \lambda + \delta \left[ 1 - \frac{f_H(y)}{f_L(y)} \right] \right\}.
$$

The limited liability constraint $s(y) \leq y$ and $\lim_{y \uparrow 10} v'(y) = \infty$ imply that there exists a value $y_1 > 0$, such that

$$
v'(s(y)) > \lambda + \delta \left[ 1 - \frac{f_H(y)}{f_L(y)} \right] \text{ for } y < y_1, \implies s^*(y) = y \text{ for } y < y_1.
$$
which establishes that there is a payoff realization $y_1$ below which all the venture’s payoff accrues to the risk averse FI. For $y > y_2$, we have $s^*(y) = y$ as well, but it should be noted that without any further assumptions on $\lim_{y \to \infty} (f_H/f_L)$, it is possible to have $y_2 = \infty$.

For $y \in (y_1, y_2)$, the limited liability constraint is not binding ($0 < s^*(y) < y$) and the optimal solution $s^*(y)$ is determined by

$$v'(s^*(y)) = \lambda + \delta \left[ 1 - \frac{f_H(y)}{f_L(y)} \right]. \quad (20)$$

Assumption 2 implies that $y^* \in (y_1, y_2)$ and for $y \in (y_1, y^*)$ the monotonicity constraint in Assumption 3 is binding. This follows, since $f_H/f_L$ is strictly decreasing for $y < y^*$ by Assumption 2, and hence $s^*(y)$ that solves (20) is strictly decreasing in that interval. Consequently, under Assumption 3 we have $s^*(y) = y_1$ for $y \in (y_1, y^*]$. Equation (20) also implies that $s^*(y)$ is strictly increasing for $y \in (y^*, y_2)$, since $f_H/f_L$ is strictly increasing for $y > y^*$ by Assumption 2. The figure below illustrates the optimal solution to (P2) under Assumptions (1-3) when $y_2 = \infty$.

![Figure 2. An optimal solution to (P2) under Assumptions (1-3)](image)

While the shape of the optimal solution in the above figure resembles that of a convertible debt contract, one should be cautious with such an interpretation for three reasons. First, unlike an equity contract the portion for $y > y^*$ is not linear. Second, we may have $s^*(y) = y$ at the upper end again, unless $y_2 = \infty$. Third, the behavior of $s^*(y)$ for $y > y_1$ is driven by the regularity condition on $f_H/f_L$ imposed by Assumption 2 and the flat portion arises due to the monotonicity restriction imposed by Assumption 3. At the same time, the analysis in this section does illustrate that under the above two additional assumptions, an optimal solution $s^*(y)$ can exhibit the two properties of convertible debt that make it superior to pure debt, pure equity and mixed debt-equity in (P2): at the lower end, the solution $s^*(y)$ assigns the whole payoff to FI for better risk sharing; and at the upper end it has a convex part that eliminates the entrepreneur’s preference for the high risk action.
5  Conclusion

This paper considers a financial contracting problem between a risk neutral entrepreneur who seeks funds for his venture and a risk averse financier who can provide financing. The riskiness of the venture’s payoff depends on the action that the entrepreneur takes after the financing is agreement. When the action choice is enforceable, a pure debt contract achieves optimal risk sharing between parties under limited liability and the financier prefers to enforce the low risk action. When the action choice is not enforceable, due to the well known risk shifting problem a debt contract induces the entrepreneur a preference for the high risk action due to its convex residual payoff. Accordingly, implementing the low risk action requires a deviation from the optimal risk sharing arrangement provided by the debt contract. I focus on situations where despite this deviation from optimal risk sharing and associated agency cost, implementing the low risk action makes the risk averse financier better off compared to opting for the high risk action. In this setting, I show that a convertible debt contract outperforms pure debt, pure equity and any mixture of debt and equity.

This result follows because a convertible debt contract has two desirable properties: (i) at the lower end of payoff realizations it assigns the whole payoff to the risk averse FI and improves risk sharing, (ii) at the upper end of payoff realizations, it creates convexity in FT’s payoff schedule and corrects EN’s high risk incentives. This role can not be achieved by simple mixtures of debt and equity, since mixed debt-equity contracts also yield a concave payoff for the financier (and hence a convex one for the entrepreneur) and implement the high risk action. For further insight, in an extension I also analyze whether an optimal solution can exhibit these two properties of convertible debt that make it superior to pure debt, pure equity and mixed debt-equity. I show that these two properties can emerge provided that (i) the admissible contracts exhibit a monotonicity property, i.e., the financier’s payoff is constrained to be non-decreasing in the venture’s payoff, and (ii) the likelihood ratio of the density functions for the high and low risk actions is first decreasing and then increasing.
Appendix

This Appendix presents an example to illustrate that FI can be better off from implementing $a_L$ even with an equity contract $s^E(y) = \pi y$ (which satisfies (14b) but not necessarily an optimal solution to (P2)) rather than opting for $a_H$ by offering $s_H(y) = \text{Min}(y, m_H)$. Specify FI’s utility function by $v(y) = y^{1-\gamma}/(1 - \gamma)$ where $\gamma \in (0, 1)$ and let

$$F_L(y) = \frac{y}{2} \text{ for } y \in [0, 2],$$

$$F_H(y) = \frac{\varepsilon}{2} + \left(\frac{1 - \varepsilon}{2}\right)y \text{ for } y \in [0, 2) \text{ and } f_H(2) = \frac{\varepsilon}{2}.$$ 

The parameter $\varepsilon \in (0, 1]$ measures the extent that $a_H$ shifts probability mass to the tails. One can verify that $F_H$ is a mean preserving spread of $F_L$ and both distributions yield an expected value $\mu = 1$. Let us set EN’s reservation wage as $w \in (0, 1)$, so that there is an expected surplus to be shared. First consider (6b) that describes the face value of the debt claim $m_H$ when FI offers $s_H(y) = \text{Min}(y, m_H)$ and implements $a_H$. Solving (6b) under the above specification yields

$$m_H = \begin{cases} 
\frac{2 - \varepsilon - \sqrt{\varepsilon^2 + 4w(1 - \varepsilon)}}{2(1 - \varepsilon)} & \text{if } \varepsilon \in (0, 1), \\
\varepsilon & \text{if } \varepsilon = 1.
\end{cases}$$ 

Consider now an equity contract $s^E(y) = \pi y$ that solves EN’s participation constraint (14a) as an equality. Under the above specification of $F_L(y)$, we have $\pi = 1 - w$. This equity contract satisfies (14b) and implements $a_L$. Denote FI’s expected utility from $s^E(y)$ by $V_L(s^E(y))$ and her expected utility from $s_H(y)$ by $V_H(s_H(y))$. When $\varepsilon = 1$, for any $\gamma \in (0, 1)$ and $w \in (0, 1)$, we have

$$V_L(s^E(y)) = \frac{2^{1-\gamma}(1-w)^{1-\gamma}}{(2-\gamma)(1-\gamma)} > V_H(s_H(y)) = \frac{2^{1-\gamma}(1-w)^{1-\gamma}}{2(1-\gamma)}.$$ 

To obtain further examples, let us set $w = 0.25$ and $\gamma = 0.5$. The table below reports $V_L(s^E(y))$ and $V_H(s_H(y))$ for $\varepsilon \in \{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}\}$. Note that in each case, the expected utility $V_L(s^E(y))$ under the equity contract $s^E(y) = \pi y$ is higher than the expected utility $V_H(s_H(y))$ under the debt contract $s_H(y)$ that implements $a_H$.

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$m_H$</th>
<th>$V_H(s_H(y))$</th>
<th>$V_L(s^E(y))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.125</td>
<td>1.06</td>
<td>1.60</td>
<td>1.64</td>
</tr>
<tr>
<td>0.250</td>
<td>1.13</td>
<td>1.56</td>
<td>1.64</td>
</tr>
<tr>
<td>0.500</td>
<td>1.26</td>
<td>1.45</td>
<td>1.64</td>
</tr>
</tbody>
</table>
References


