Knightian Uncertainty and

Capital Structure


by


Seokwoo Lee


A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in the University of Michigan
2013

*Doctoral Committee:*

    Associate Professor Uday Rajan, Chair
    Assistant Professor Ing-Haw Cheng
    Associate Professor Venkatesh Nagar
    Associate Professor Amiyatosh Purnanandam
    Professor Sijue Wu

# Dedication

To my mother and father
and to my wife Jungyun
and to my child Julia

# Acknowledgements

# Contents

# List of Figures

# List of Tables

# Abstract

My dissertation aims to understand a firm's optimal capital structure decision when it confronts Knightian uncertainty. It contains three chapters.

In Chapter One, I derive the optimal capital structure of a firm when its manager is ambiguity-averse. My model predicts substantially lower leverage for such firms, in comparison to traditional static trade-off models. I use the 1982 Voluntary Restraint Agreement (VRA) on steel import quotas between the U.S. government and the European Community as an exogenous reduction in Knightian uncertainty faced by firms in the U.S. steel industry. Using a difference-in-difference methodology, I find that when uncertainty is resolved, a median firm in the U.S. steel industry increases its market and book leverage by approximately 12% relative to a matched control firm from another industry. The results are not explained away by changes in traditional risk factors or by a change in expected future profitability.

In Chapter Two, I develop a new measure of Knightian uncertainty using a mixture of normal distributions. Using this novel measure, I find that a median firm in the sample decreases its market leverage by approximately 3.6%, when the firm's measure of uncertainty increases by one standard deviation. I also find an increase in a firm's uncertainty measure from the minimum to the maximum is associated with a 6.3% increase in the firm's propensity to take almost zero leverage.

In Chapter Three, I consider a firm's decision to exercise real options, when the firm faces Knightian uncertainty about the evolution of project value. I provide two continuous time applications extended from the discrete time applications in Miao and Wang (2011) that show the value of waiting (which increases with risk) decreases with uncertainty. I incorporate ambiguity-averse investors who are uncertain about the evolution of a firm's assets into the dynamic capital structure model of Leland (1994). I show that even in this dynamic setting, uncertainty provides a more plausible explanation for firms taking on low leverage than risk alone.

# Chapter 1

# Knightian Uncertainty and Capital Structure: Theory and Evidence

## 1.1 Introduction

> "··· there is no possibility of forming in any way groups of instances of sufficient homogeneity to make possible a quantitative determination of true probability. Business decisions, for example, deal with situations which are far too unique, generally speaking, for any sort of statistical tabulation to have any value for guidance. The concept of objectively measurable probability or chance is simply inapplicable."
>
> – (Knight, 1921, p. 231)

According to Knight (1921), risk applies to situations where agents can accurately measure the odds of prospects. Uncertainty (or ambiguity), on the other hand, applies to situations where agents do not have the information necessary to assess probabilities in the first place. For example, an automobile company might forecast that the risk of producing a defective car is exactly one out of 2,000, yet they will not be able to accurately forecast the economic outlook for the automobile industry in 30 years because there are too many unknown factors to calculate. Hence, uncertainty in this sense is common in business decisions.

To financial economists, it is a puzzle that firms take on lower leverage than predicted by trade-off models of capital structure (Miller, 1977; Bradley et al., 1984; Graham, 2000). The median corporate market leverage ratio in the United States between 1990 and 2008 is

around 20% (Minton and Wruck, 2001; Lemmon and Zender, 2001). The traditional models based on pure risk (not Knightian uncertainty) often predict leverage ratios in excess of 50%, given reasonable exogenous parameters (Berk and DeMarzo, 2008; Frank and Goyal, 2008). Moreover, Strebulaev and Yang (2012) document a zero-leverage puzzle: between 1962 and 2006, about 10% of public non-financial firms take on zero debt, and 32% have zero or negative net debt.

To address these puzzles, I examine the effect of Knightian uncertainty on leverage. I incorporate a manager's ambiguity aversion into a static trade-off model. I derive the optimal capital structure when the manager faces Knightian uncertainty over the firm's future cash flow. The model shows that the amount of uncertainty and the manager's aversion toward it lead to substantially lower leverage than the traditional counterpart. The intuition is clear. A Knightian manager is unsure about the true distribution of the next period's cash flow. Instead, she maintains a set of distributions (priors) that she thinks are plausible, and displays aversion toward this uncertainty. Her ambiguity aversion causes her to pay more attention to the outcomes of relatively pessimistic priors. Hence, the marginal cost of default increases with both the amount of uncertainty perceived by the manager and her aversion toward it, while the tax benefit of debt decreases. To balance this trade-off, the ambiguity-averse manager takes on considerably lower leverage than she would in the absence of uncertainty.

I show that there is a distinction between the effect of uncertainty and the effect of risk on leverage. A firm's optimal debt does not strictly decrease with risk. As in Bradley et al. (1984), when the firm's cash flow is highly volatile, the manager may want to increase debt to increase the value of tax shields. I show that it is true even when the manger is risk-averse. In contrast, in the presence of ambiguity, an ambiguity-averse but risk-neutral manager strictly decreases debt as the level of ambiguity rises. In particular, I show that an ambiguity-averse but risk-neutral manager takes on strictly less debt than an ambiguity-neutral but risk-averse manager, given a fixed level of risk. Therefore, uncertainty provides a more plausible explanation for firms having low leverage than risk alone.

To empirically test this prediction, I use a difference-in-difference methodology to estimate the causal effect of the resolution of uncertainty on leverage. I consider the 1982 Voluntary Restraint Agreement (VRA) on steel import quotas between the U.S. government and the European Community (EC) as an exogenous reduction in the amount of Knightian uncertainty faced by U.S. steel firms. Prior to this agreement, U.S. steel manufacturers faced considerable Knightian uncertainty over the outcomes of antidumping (AD) and countervailing duty (CVD) legal proceedings that they filed against European steel producers. The 1982

2

VRA resolved the uncertainty confronted by the U.S. steel industry in the pre-VRA period, providing an empirical setting to estimate the causal effect of the resolution of uncertainty on leverage. The key identification assumption is that the 1982 VRA is exogenous to the leverage decision of U.S. steel manufacturers. The null hypothesis is that when uncertainty is resolved by the 1982 VRA, the average change in leverage is the same for both U.S. steel firms (the treatment) and the matched control firms. The results suggest that when uncertainty is resolved by the 1982 VRA, a median firm in the steel industry increases its market and book leverage by 11.5% and 12.3%, respectively, relative to a matched control firm from another industry. These results are economically strong and statistically significant. More importantly, the results are not explained away by changes in traditional factors such as risk, bankruptcy cost, and future profitability.

I conduct a few placebo tests in which I assume the VRA happens at different times, instead of in 1982. Steel firms and matched control firms respond similarly to the placebo shocks. In addition, following Bertrand et al. (2004), I time-average the data before and after the 1982 VRA shock and run a difference-in-difference analysis on this averaged outcome variable. My results remain robust.

One potential concern is that the U.S. steel firms may have lobbied for the enactment of the 1982 VRA. If leverage (the outcome variable) is a direct reason for the passage of the VRA, this poses a threat to the internal validity of the difference-in-difference methodology. However, any lobbying by steel firms is likely to be driven by concerns over profitability rather than issues related to leverage alone. Then, as long as the profitability of a firm is included in the regression as a firm-specific control variable, the difference-in-difference estimator identifies the causal effect of the resolution of uncertainty on a steel firm's leverage. Moreover, the U.S. government independently had a strong incentive to negotiate the VRA with the EC, even absent any lobbying by steel firms. The U.S. government had a concurrent dispute with the EC over a natural gas pipeline from the Soviet Union to Western Europe (the Trans-Siberian Pipeline). The U.S. government believed that the punitive duties on steel imports would make talks over the pipeline issue even more problematic and impede cooperation with Europe (Moore, 1996). During the Cold War (1947-1991), the U.S. government saw it as a critical security issue (The Weinberger Paper, Pentagon Report, 1981). Therefore, I argue that the 1982 VRA may be treated as exogenous to the leverage of U.S. steel firms.

There is a long-standing body of research on explaining the low-leverage puzzle. A number of recent dynamic trade-off models have been able to predict lower optimal leverage ratios (Goldstein et al., 2001; Morellec, 2004; Strebulaev, 2007). The dynamic trade-off models that incorporate endogenous investment are also able to predict lower optimal leverage

ratios (Hennessy and Whited, 2005, 2007; DeAngelo et al., 2011). These models represent an important step toward explaining the low-leverage puzzle. My contribution is to propose a simple model that provides a plausible new reason (a manager's ambiguity-aversion) for firms having low leverage. A second contribution of this paper is to present empirical evidence that suggests there is a causal relationship between uncertainty and leverage, adding to what we know of the empirical determinants of leverage (e.g., Titman and Wessels, 1988; Rajan and Zingales, 1995; Lemmon et al., 2008). My results suggest that the amount of uncertainty perceived by a firm is an important determinant of leverage.[1]

The Ellsberg Paradox (Ellsberg, 1961) illustrates the Knightian distinction between risk and ambiguity. In experiments, agents are typically ambiguity-averse: they prefer a risky bet to an uncertain one.[2] Gilboa and Schmeidler (1989) show that the Ellsberg paradox can be resolved by using multiple priors and a maxmin objective. An ambiguity-averse agent computes expected utility under the worst-case scenario among the priors that he thinks are possible. Klibanoff et al. (2005; henceforth, KMM) extend the multiple priors model to allow smooth preferences toward ambiguity. Their model provides a separation between the attitude to ambiguity and the amount of the ambiguity itself. In this paper, I use the KMM framework to model the amount of ambiguity perceived by a manager and her aversion toward it.

In Section 2, I begin by providing an outline of the smooth ambiguity aversion model proposed by KMM. Then, I consider a simple static trade-off model of capital structure and incorporate the manager's ambiguity aversion into the model to derive the firm's optimal capital structure. I also consider the relevant comparative statics of the ambiguity-aversion augmented model. In Section 3, I describe the nature and the timing of 1982 VRA, provide the details of the difference-in-difference methodology, as well as the construction of the matched sample. The main empirical results are presented in Section 4. In Section 5, I perform robustness checks for the results. Section 6 concludes.

---

[1]In Chapter 2, I propose a new measure of Knightian uncertainty and find that there is an economically sizeable and statistically significant negative relationship between uncertainty and leverage.

[2]Details about the Ellsberg Paradox are in Appendix 1.A.

## 1.2 Theory: Trade-off Model with a Manager's Ambiguity Aversion

### 1.2.1 Smooth ambiguity aversion preference

This section provides a brief review of Klibanoff, Marinacci and Mukerji (2005)'s smooth ambiguity aversion model. For clear exposition, I introduce a discrete version of the KMM model. Consider an agent who cares about the realization of a random variable $X$. The agent is uncertain about the true distribution of $X$. He maintains $\Pi = \{F_1, \cdots, F_n\}$ as the set of distributions that he thinks are plausible as the true distribution for $X$. It is evident that if $\Pi$ is a singleton, then there is no ambiguity about the distribution of $X$. The agent attaches his own subjective belief $\mu_i$ to $F_i$ in $\Pi$, being the true distribution of the random variable $X$.

The smooth ambiguity aversion preference is expressed as follows: First, the agent evaluates all possible first-stage expected utilities of $X$ corresponding to a distribution $F_i$ in $\Pi$. He gets a set of the (first-stage) expected utilities, each being indexed by $i$.

$$\text{Step 1:} \qquad \int u(x) dF_i \qquad \text{each } F_i \in \Pi$$

where $u(\cdot)$ is a usual Bernoulli utility function. Second, the agent transforms each first-stage expected utility by an increasing function $\phi(\cdot)$.

$$\text{Step 2:} \qquad \phi\left(\int u(x) dF_i\right) \qquad \text{each } F_i \in \Pi$$

Finally, the agent obtains the second-stage expectation $V(x)$ by averaging the transformed first-stage expected utilities by his subjective belief $\mu_i$ attached to a corresponding $F_i$ in $\Pi$ for $i = \{1, \cdots, n\}$.

$$\text{Step 3:} \qquad V(x) := \sum_{i=1}^{n} \mu_i \cdot \phi\left(\int u(x) dF_i\right) \tag{1.1}$$

In this representation, the agent's preference is "an expected utility over expected utilities."

The role of the function $\phi(\cdot)$ in the second step is critical. The shape of $\phi(\cdot)$ describes the agent's attitude towards ambiguity, analogous to the attitude toward risk characterized by the shape of utility function $u(\cdot)$. In particular, if $\phi(\cdot)$ is concave, the agent evaluates the final stage expectation of the concave transformed first-stage expected utilities. As with risk

aversion, the concave transformation reflects his aversion to ambiguity; he attaches a larger weight to lower first-stage expected utility. If $\phi$ is convex, as with risk seeking, the agent tends to put a higher weight on higher first stage expected utility. Accordingly, he displays ambiguity-seeking behavior. On the other hand, the amount of ambiguity is characterized by the dispersion of the set of priors $\Pi$. KMM emphasize that such separation is a distinction of their model from other models, including the multiple priors model.[3]

## 1.2.2 Trade-off model with ambiguity aversion

### 1.2.2.1 The model

The model is based on a standard trade-off model (Kraus and Litzenberger, 1973; Bradley et al., 1984). The firm's leverage is determined by a single period trade-off between the tax benefit of debt and the deadweight costs of bankruptcy. Investors are assumed to be risk and ambiguity neutral. The firm faces a constant marginal tax rate on end-of-period cash flow. If the firm is unable to make the promised debt payments, it incurs deadweight financial distress costs. If the cash flow is negative, due to limited liability, no debt is paid. If the cash flow is positive but not enough to cover the debt payment $B$, the firm defaults. There is a deadweight loss of $kX$ in the process. When the cash flow is enough for the firm not to default, equity holders receive $(X - B)(1 - \tau)$.

Suppose the firm's next period cash flow $(X)$ has a distribution $F(x)$. Then, the market value of debt is:

$$D(B) = \left(\frac{1}{1+r}\right)\left(\int_0^B x(1-k)\,dF(x) + \int_B^\infty B\,dF(x)\right) \qquad (1.2)$$

and the market value of equity is:

$$E(B) = \left(\frac{1}{1+r}\right)\left(\int_B^\infty (1-\tau)(x-B)\,dF(x)\right), \qquad (1.3)$$

where $r$ is a risk-free rate. The levered firm value is $V(B) = D(B) + E(B)$. It is useful to

---

[3]In the max-min model, the manager always chooses the worst-case prior in $\Pi$. Hence, the dispersion of $\Pi$ represents both the amount of ambiguity and the attitude toward it. If the manager is infinitely ambiguity-averse, the smooth ambiguity model in effect coincides with the multiple priors model (Klibanoff et al., 2005).

rewrite $V(B)$ as:

$$V(B) = \frac{1}{1+r} \left( \underbrace{\int_0^\infty (1-\tau)x\, dF(x)}_{\text{unlevered firm value}} - \underbrace{\int_0^B k\, x\, dF(x)}_{\substack{\text{deadweight loss} \\ \text{due to bankruptcy}}} + \underbrace{\tau \int_0^B x\, dF(x) + \tau \int_B^\infty B\, dF(x)}_{\text{Tax benefit from debt}} \right).$$

I assume that a manager faces Knightian uncertainty and is averse toward it. However, I assume that she is still risk-neutral. As a Knightian decision maker, the manager is uncertain about the true distribution of cash flow, $F(x)$. Instead, she maintains $\Pi$, a set of all possible distributions $\{F_i(x)\}_{i=1}^n$ that she thinks are plausible as the true distribution for the cash flow. She associates the subjective belief $\mu_i$ to each $F_i(x)$ in $\Pi$. When making a decision about the optimal leverage, the manager considers the expected firm value for each $F_i$ and displays aversion to ambiguity characterized by the concave function $\phi(\cdot)$. Because of her ambiguity aversion, she assigns higher weights to firm values with relatively pessimistic priors in $\Pi$. Therefore, she makes a lower leverage decision than in the absence of ambiguity. When $\Pi$ is a singleton, the manager is absolutely sure about the true distribution of cash flow, thereby the model becomes a traditional single-period static trade off model.

Precisely, the manager chooses the optimal leverage to maximize the firm value applying the smooth ambiguity-averse preference:

$$\max_{B>0} \underbrace{\sum_{i=1}^n \phi(V(B|F_i)) \cdot \mu_i}_{=\mathbb{E}^\mu\left[\phi(V(B|F))\right]}. \tag{1.4}$$

Here, I slightly abuse the notation. $V(B|F_i)$ is the levered firm value when it takes the debt amount $B$ given a distribution $F_i(x)$ in $\Pi$.

The first-order condition is:

$$\mathbb{E}^\mu\left[\phi'(V(B|F))\frac{\partial V(B|F)}{\partial B}\right] = 0. \tag{1.5}$$

As usual, the optimal debt level $B$ can be obtained by solving (1.5) for $B$.

Lemma 1.1 states that when the ambiguity-averse manager chooses the optimal debt, she behaves as if she were ambiguity-neutral but pays more attention to relatively pessimistic priors. It is a special case of Proposition 5 of Rigotti et al. (2008).

**Lemma 1.1** (Existence). *The first-order condition* (1.5) *is equivalent to:*

$$\mathbb{E}^{\mu}\left[\phi'(V(B|F))\frac{\partial}{\partial B}V(B|F)\right] = \mathbb{E}^{\mu^*}\left[\frac{\partial V(B|F)}{\partial B}\right] = 0. \tag{1.6}$$

*Formally, given her subjective belief $\mu$ to the set $\Pi$ of priors, there exists a Radon-Nykodym derivative $Z$ such that:*

$$d\mu^* = Z \cdot d\mu\,, \tag{1.7}$$

*where:*

$$Z = \frac{\phi'(V(B|F))}{\mathbb{E}^{\mu}(\phi'(V(B|F)))} \qquad and \qquad \mathbb{E}^{\mu}Z = 1\,.$$

*Proof.* See Appendix 1.B.1. □

The belief $\mu^*$ is interpreted as an adjusted belief: the subjective belief $\mu$ is adjusted by $\phi'(V(B|F))$. I refer to $\mu^*$ as an ambiguity-adjusted belief in the following sense. In a case where the manager's subjective belief $\mu$ on $\Pi$ is assumed to be uniform (*i.e.*, she thinks all distributions in $\Pi$ are equally likely), her ambiguity aversion forces her to put the largest weight on the worst case prior that yields the highest marginal value $\phi'(\cdot)$. Note that Lemma 1.1 states the *existence*, not the *uniqueness* of a density $Z$. I thus obtain the following corollary to Lemma 1.1.

**Corollary 1.1.** *The optimality condition* (1.6) *becomes:*

$$kB \cdot \mathbb{E}^{\mu^*}[f(B)] = \tau\left(1 - \mathbb{E}^{\mu^*}[F(B)]\right), \tag{1.8}$$

*where $\mathbb{E}^{\mu^*}[f(x)] = \sum_{i=1}^{n}\mu_i^* \cdot f_i(x)$, a mixture of densities $f_i(x)$ with ambiguity adjusted belief $\mu_i^*$.*[4]

*Proof.* See Appendix 1.B.2. □

The intuition of Corollary 1.1 is as follows. The Knightian manager is uncertain about the true density $(f(B))$. Instead, she finds the set of densities that she thinks are plausible $(\{f_i(B)\}_{i=1}^{n})$ and averages them with her ambiguity-adjusted belief $(\mathbb{E}^{\mu^*}[f(B)])$. Similarly,

---

[4]When the manager is absolutely sure about the distribution $F(x)$ (i.e., the set $\Pi$ is a singleton), then the optimality condition (1.8) becomes the usual first-order condition:

$$\frac{\partial V(B)}{\partial B} = kBf(B) - \tau(1 - F(B)) = 0,$$

because when $\Pi$ is a singleton, $\mu$ becomes degenerate, therefore $f^*(x)$ is $f(x)$.

she also maintains the set of distributions of the firm being solvent ($\{\int_B^\infty f_i(x)dx\}_{i=1}^n = \{1 - F_i(B)\}_{i=1}^n$), then obtains the ambiguity-adjusted distribution of the firm being solvent ($1 - \mathbb{E}^{\mu^*}[F(B)]$). The optimality condition (1.8) means that the ambiguity-adjusted marginal cost of default and the ambiguity-adjusted marginal tax benefit of debt must be equal.

### 1.2.2.2 Comparative statics of the optimal leverage

Now, I determine two comparative statics of the optimal leverage with respect to the amount of uncertainty perceived by the manager and degree of her aversion toward it. Proposition 1.1 states that (a) while keeping the amount of uncertainty fixed, as the manager's ambiguity aversion rises, she may reduce the optimal level of debt, and (b) while keeping the manager's ambiguity aversion fixed, as the amount of uncertainty she faces increases, she may reduce the optimal level of debt.

The amount of uncertainty perceived by the manager is measured by the dispersion of $\Pi$. I denote it as $\delta$. For this section, I assume that all distributions in $\Pi$ are gaussian and variance of them is common and known, therefore the manager is unsure only about the mean of the distribution for the firm's cash flow. Then, the dispersion of $\Pi$ is defined as $\delta = \max |\theta_i - \theta_j|$ for all $i \neq j$, where $\theta_i$ is the mean parameter of distribution $F_i$ in $\Pi$.[5]

Meanwhile, the degree of a manager's ambiguity aversion is characterized by the degree of concavity of $\phi(\cdot)$. Following Klibanoff et al. (2005), I assume that the manager displays constant absolute ambiguity aversion (CAAA). That is,

$$\phi(x) = \begin{cases} 1 - e^{-\alpha x} & \alpha > 0 \\ x & \alpha = 0 \end{cases} \tag{1.9}$$

where the parameter $\alpha$ is the coefficient of absolute ambiguity aversion such that:

$$\alpha = -\frac{\phi''(x)}{\phi'(x)}.$$

Similar to the coefficient of risk aversion, $\alpha$ represents the degree of concavity of $\phi(\cdot)$, and therefore represents the degree of the manager's ambiguity aversion.

**Proposition 1.1.** *Ceteris paribus, the optimal amount of debt B that satisfies the first-order*

---

[5]In general, the dispersion of $\Pi$ can be characterized by the maximum statistical distance $\mathcal{D}(\cdot||\cdot)$ between any two distributions in $\Pi$, $\max \mathcal{D}(F_i||F_j)$ for all $i \neq j$ in $\{1, \cdots, n\}$.

*condition (1.8) decreases with α, the manager's degree of ambiguity aversion. That is,*

$$\frac{\partial B}{\partial \alpha} < 0 \,.$$

*Also, the optimal amount of debt B decreases with the amount of uncertainty that the manager perceives. Let the amount of uncertainty be δ. Then,*

$$\frac{\partial B}{\partial \delta} < 0 \,.$$

*Proof.* See Appendix 1.B.3 on page 28. □

The intuition is as follows. By Lemma 1.1 (therefore, Corollary 1.1), as ambiguity-aversion ($\alpha$) increases, the manager puts higher weight on relatively pessimistic priors in $\Pi$. Hence, given an optimal debt level $B$, as $\alpha$ increases, the ambiguity-adjusted marginal cost of default ($kB \cdot \mathbb{E}^{\mu^*}[f(B)]$) increases, while the ambiguity-adjusted marginal tax benefit of debt ($\tau \left(1 - \mathbb{E}^{\mu^*}[F(B)]\right)$) decreases. To balance this trade-off, the manager must optimally reduce the amount of debt.

The optimal leverage ratio is defined as:

$$\text{Leverage ratio} = \frac{B^*}{E(B^*) + B^*}, \tag{1.10}$$

where $B^*$ is the optimal amount of debt computed from the first-order condition (1.8), and $E(B^*)$ is the market value of equity at $B^*$ using (1.3). Figure 1.1 on page 38 illustrates a numerical example of the comparative statics of the optimal leverage ratio with respect to the variation in the manager's degree of ambiguity aversion. The computation details are in Appendix 1.C.

Next, as the amount of uncertainty ($\delta$) increases, the pessimistic prior in $\Pi$ gets worse (i.e., shifted left). According to the optimality condition (1.8), the ambiguity adjusted marginal cost of default ($kB\mathbb{E}^{\mu^*}[f(B)]$) increases with $\delta$, while the ambiguity-adjusted marginal tax benefit of debt decreases. Therefore, the ambiguity-averse manager must optimally reduce the optimal level of debt to balance this trade-off. Figure 1.2 illustrates a numerical example of the comparative statics of the optimal leverage ratio with respect to the variation in the amount of uncertainty. The leverage ratio at the given optimal debt $B^*$ is computed using (1.10). The computation details are in Appendix 1.C.

From Proposition 1.1 (see also Figures 1.1 and 1.2 on page 38), the effect of uncertainty on leverage is clear. As uncertainty increases, leverage strictly decreases. One can ask whether

risk aversion without invoking ambiguity aversion is also sufficient. Qualitatively, the effects of uncertainty and risk on leverage are similar. Because the certainty equivalent of the risk-averse manager decreases with risk, the risk-averse manager also tends to employ lower leverage than the risk-neutral manager. Quantitatively, however, the effect of uncertainty on leverage is higher than that of risk. Moreover, the risk-averse manager's choice of leverage does not strictly decrease with the risk of cash flow, whereas the ambiguity-averse manager's choice of leverage strictly decreases with uncertainty.

**Proposition 1.2.** *For a risk-averse but ambiguity-neutral manager, there exists a level of risk, $\overline{\sigma}$, such that (i) the optimal debt, B, strictly decreases with risk when the current risk is lower than $\overline{\sigma}$,*

$$\frac{\partial B}{\partial \sigma} < 0 \qquad if \qquad \sigma < \overline{\sigma}$$

*(ii) the optimal debt, B, increases or remains the same with increasing risk when the current risk is higher than $\overline{\sigma}$,*

$$\frac{\partial B}{\partial \sigma} \geq 0 \qquad if \qquad \sigma \geq \overline{\sigma}$$

*Proof.* See Appendix 1.B.4. $\overline{\sigma}$ is a solution of the integral equation (1.20) on page 30 in Appendix 1.B.4. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

In contrast to Proposition 1.1, Proposition 1.2 states that the optimal level of debt does not strictly decrease with risk. When a firm's cash flow is less volatile, an increase in risk sufficiently increases the marginal cost of default, but decreases the marginal tax benefit at the current optimal level of debt. Therefore, she must decrease the optimal level of debt to balance this trade-off. When the cash flow of a firm is highly volatile, the manager is more likely to gamble on the chance of the firm being solvent in the next period. That is, when the present risk is high, the default intensity can decrease with risk. But, the manager can gain additional tax benefit of debt by adding an additional amount of debt, as long as the firm is still solvent in the next period. Therefore, when the firm's cash flow is highly volatile, the manager may want to increase debt to increase the value of tax shields. Bradley et al. (1984) also point out that when the manager of a firm is risk-neutral, the effect of risk on the optimal level of debt is undetermined. In Proposition 1.2, I show that it is true even when the manager is risk-averse.

Figure 1.3 on page 39 illustrates numerical examples of Propositions 1.2. The leverage ratio at the risk-averse but ambiguity-neutral manager's optimal debt choice $B^*$ is computed using (1.10). The figure depicts that the risk-averse but ambiguity-neutral manager's optimal leverage does not strictly decrease with risk after a certain level of risk with the reasonable

proportional deadweight cost ($k = 0.4$) and marginal tax rate ($\tau = 0.35$).

Next, consider two managers. Manager "A" is risk-neutral but ambiguity-averse with CAAA preference. Manager "R" is risk-averse with CARA preference, but is ambiguity-neutral. For a given level of risk, the optimal leverage ratio of manager "A", is strictly less than that of manager "R," if and only if, the amount of uncertainty perceived by the managers ($\delta$), the level of the risk of the firm's cash flow ($\sigma$), the coefficient of absolute ambiguity aversion ($\alpha$) of the manager "A", and the coefficient of absolute risk aversion ($\rho$) of the manager "R" satisfy the following condition:

$$\rho((1 - \tau)x_0 + \tau B_R - B_0) < \alpha(V(B_R|\theta_2) - V(B_R|\theta_1)), \tag{1.11}$$

where $B_0 \in [(1 - k)B, B]$ satisfies

$$kB\, u'(B_0)\, f(B) = \int_B^\infty u'((1 - \tau)x + \tau B)f(x)dx, \tag{1.12}$$

and the optimal leverage taken by the manager "R", $B_R$, satisfies

$$-KB_R u'(B_0) \int_\Pi f(B_R|\theta)d\mu(\theta) + \tau \int_\Pi \int_{B_R}^{\bar{x}} u'((1 - \tau)x + \tau B_R)f(x|\theta)dxd\mu(\theta) = 0. \tag{1.13}$$

Note that $\Pi = \{f(\cdot|\theta)\}_{\theta \in \Theta}$ is the compact set of prior about the cash flow distribution, $\mu$ is associated belief over $\Pi$, $u(\cdot)$ is the CARA function, and $\phi(\cdot)$ is the CAAA function given by (1.9) on page 9. The details to derive the condition (1.11) are available in Appendix 1.B.5 on page 31.

Because condition (1.11) is less intuitive, I refer to Figure 1.3 on page 39 to provide the quantitative effect of uncertainty on leverage as compared to that of risk on leverage. When the coefficient of ambiguity aversion ($\alpha = 3$) is comparable to that of risk aversion ($\rho = 2$) and the amount of uncertainty is fixed as $\delta/\sigma = |\theta_{\min} - \theta_{max}|/\sigma = 2$, the manager "A" takes on strictly less leverage than the manager "R", given a fixed level of risk. The computation details to generate the figure are in Appendix 1.D on page 35.

The comparability between the coefficient of ambiguity aversion $\alpha = 4$ and that of risk aversion $\rho = 2$ has to be justified. Consider the following thought experiment related to the Ellsberg paradox following Ju and Miao (2007). The idea is that the ambiguity premium computed as the difference between the certainty equivalent from risky situation and non-ambiguity equivalents from ambiguous situation must be reasonable. Camerer (1999) reported that the ambiguity premium is typically on the order of 10–20 percent of the ex-

pected value of a bet in the Ellsberg-style experiments (Ju and Miao, 2007). Given this evidence, the choice of ambiguity aversion parameter ($\alpha = 4$) in Figure 1.3 on page 39 seems reasonable with respect to the risk aversion parameter ($\rho = 2$). The calibration details are available in Appendix 1.A on page 26.

Strebulaev and Yang (2012) document the puzzling evidence that the substantial number of large public non-financial U.S. firms follow a zero-debt or almost zero-debt policy. According to Proposition 1.1 and 1.2, uncertainty has a better chance to explain the tendency of firms taking extremely low or zero leverage because the optimal leverage strictly decreases with uncertainty, whereas it does not strictly decrease with risk. In Chapter 2, I report an empirical result supporting this theoretical prediction: the amount of uncertainty perceived by a manager is positively related to the firm's propensity of taking almost zero-leverage.

## 1.3   Evidence: Resolution of Uncertainty on Leverage

For an empirical test, I use difference-in-difference methodology to estimate the causal effect of the resolution of uncertainty on leverage. I consider the 1982 VRA on steel import quotas between the U.S. government and the EC as an exogenous reduction in the amount of Knightian uncertainty faced by U.S. steel firms. Prior to this agreement, U.S. steel manufacturers faced considerable Knightian uncertainty over the outcomes of AD and CVD legal proceedings that they filed against European steel producers.

**Prediction.** *By Proposition 1.1, when uncertainty perceived by U.S. steel firms is resolved by the 1982 VRA, steel firms increase leverage relative to (matched) control firms in the post-VRA period.*

In the next section, I provide the details of the nature and the timing of the 1982 VRA.

### 1.3.1   1982 VRA impact on U.S. steel imports

In January 1982, U.S. steel companies filed a large number of AD and CVD petitions against European steel producers due to a substantial increase in imports from the EC (February 14, 1982, *The New York Times*). According to the World Bank's AD/CVD database, on January 11, 1982, major U.S. steel manufacturers including Bethlehem Steel, U.S. Steel, Republic Steel, Inland Steel, Jones and Laughlin Steel, National Steel, and Cyclops Steel filed 33 AD and 61 CVD petitions against eight EC countries.

However, the AD and CVD laws were long and cumbersome for U.S steel firms. These laws require the complainant to do an enormous amount of investigation simply to file a complaint (Range, 1980, p. 283). Once a complaint is filed, the ITA that administers the law has considerable discretion over whether it will pursue the complaint. Even if the ITA agrees to do so, the complainant must go through long and costly procedures (Solomon, 1978, Section 2). Both the formulation and investigation of a complaint are likely to be impeded by the foreign producers reluctant to provide sensitive data to the ITA (Solomon, 1978, footnote 12). And relief, if granted, consists of the imposition of a duty rather than an award of damages to the complainant. Such relief does not necessarily benefit the complainant, because the importer can shift to the other foreign producers not subject to the duties (Solomon, 1978, p. 449).

Indeed, the ITA's highly discretionary interpretation of AD/CVD laws results in highly arbitrary injury determinations (Caine, 1981). Part of the reason is that the ITA often lacks the critical information to determine the fairness of an exporter's selling price.[6] Accordingly, the ITA is regularly forced to make subjective accounting determinations (Harvard Law Review, 1983; Solomon, 1978). An example of the arbitrariness of this process is provided by the wide disparity between injury determinations calculated by the ITA in its preliminary CVD proceedings and those calculated in its final proceedings in the steel cases. For instance, in June 1982 the margin of subsidy found by the ITA for the British Steel Corporation was 40.4%, but in August 1982 the margin was 20.3%. (August 26, 1982, *The New York Times*).[7] Hence, although the steel firms know that the cases will resolve, they are unsure of the likelihood of the outcomes because of the discretionary scope and arbitrariness of the decision process of the ITA. Therefore, I claim that the U.S. steel industry in the pre-VRA period was more likely to face a high degree of Knightian uncertainty.

One potential concern is that the intervention-seeking steel industry may have lobbied for the 1982 VRA. One can ask whether or not it is a potential threat to the internal validity of difference-in-difference methodology. However, I argue that the 1982 VRA is not an event driven by the U.S. steel firms' lobbying because the U.S. government had strong incentives to negotiate VRA with the EC. It wanted to avoid the open-ended and prohibitive duties on many European steel exports. Complicating matters was a concurrent dispute with the EC

---

[6]For the fair value calculation, it is necessary to make adjustments for differences in the physical characteristics of the merchandise in the markets being compared, differences in the quantities being sold, and differences in the circumstances of the sale (the credit terms, grantees, technical assistance, advertising, and services being provided by the seller). In addition, fluctuations in exchange rates must be considered.

[7]In addition, the decisions of laws are not strictly enforceable, if favorable to the complainant. As stated in *The New York Times* (August 5, 1982): "Ian MacGregor, chairman of the British Steel Corporation, insisted that if the American decision to impose stiff duties on steel imports from Common Market countries went into effect, he would take legal action to overturn it."

over a natural gas pipeline from the Soviet Union to Western Europe (the Trans-Siberian Pipeline) (Moore, 1996).[8]

During the Cold War, the U.S. government worried that the Trans-Siberian Pipeline would make the EC dangerously dependent on Russian energy sources and that the earnings flowing from Western Europe to the Soviet Union would relieve many of the Soviet's economic problems, increasing the Soviet's military strength. Increased Soviet strength was likely to significantly add to the U.S. defense burden (The Weinberger Paper, Pentagon Report, 1981).[9] To impede the growth of the Soviet's military power and economic leverage, the U.S. government strongly opposed the construction of the Trans-Siberian Pipeline. It embargoed U.S. companies' selling supplies for the pipeline's construction and asked their European allies to deny supplying parts for the pipeline. Unexpectedly, in mid-1982, Britain and France defied the Reagan Administration's sanctions, insisting that the contract between the Soviet Union and European companies had to be honored (January 11, July 9, August 3, and September 2, 1982, *The New York Times*). The U.S. government believed that punitive duties on steel exports would make talks over the pipeline issue even more problematic and impede cooperation on what it saw as a critical security issue (The Haig Paper, Department of State Report, 1981).[10]

Moreover, a VRA program provides distinct political advantages for the U.S. government over AD/CVD duties or legislative quotas.[11] A VRA program insures that the government would retain controls over the trade policy decisions. Such discretion would not have been possible if final AD and CVD duties had been imposed.[12] A VRA program also enables the government to control the timing of protection offered to industry. Unlike AD/CVD duties, it specifies an expiration date (Caine, 1981).[13] In addition, the free-market Reagan administration could assert to the public that the government did not impose the tariffs

---

[8]The Trans-Siberian Pipeline project was first proposed in 1978 and was constructed in 1982-1984. It created the transcontinental gas transportation system from the western Ukraine to Central and Western European countries (Hardt, 1982).

[9]Declassified 2004. `http://www.margaretthatcher.org/document/110933`

[10]Declassified 2004. `http://www.margaretthatcher.org/document/110934`

[11]The Reagan Administration preferred VRA to legislative quotas. VRAs were perceived as less rigid and less permanent than legislative quotas. Second, the Administration feared that, once the Congress gave favorable treatment to one industry, it would be likely to provide comparable benefits to the other industries. Third, VRA would provide a way of circumventing the General Agreement on Tariffs and Trade (GATT)'s prohibition of quotas or would at least reduce the danger of retaliation by other countries under GATT (Lowenfeld, 1983; Harvard Law Review, 1983).

[12]The ITA can impose the final duties without any direct involvement of either the president or any other elected officials (Harvard Law Review, 1983) ([see also Anderson (1993) for the empirical evidence]).

[13]Duties can remain indefinitely subject to a five-year review. Article 11 of the General Agreement on Tariffs and Trade 1994 states "··· anti-dumping duty shall remain in force only as long as and to the extent necessary to counteract dumping which is causing injury."

but negotiated the agreement, to retain their free-trade rhetoric. From the Europeans' perspective, they also had an incentive to negotiate with the U.S. According to Viscount Davignon, the Commissioner of Industry for the European Economic Community, "··· the ceiling is preferable to the duties because the duties would have reduced trade to a trickle and cost thousand of jobs in the Europe." (October 22, 1982, *The New York Times*) These factors strongly induced the U.S. to enter negotiations with the EC for the 1982 VRA (Moore, 1996).

Second, empirical evidence that the 1982 VRA was driven by the steel firms' lobby is weak. Theoretically, the steel industry has attributes consistent with successful lobbying characteristics: relatively small numbers of actors in the group and loyal unionization. Empirically, however, measures of lobbying power, such as concentration and unionization, are poor explanatory variables for obtaining the government's protection (see Trefler 1993; Markusen 1996). The empirical findings documented in Trefler (1993) refer to general government protections rather than VRAs, in particular. However, whether the steel industry's lobbying was effective for the 1982 VRA is empirically questionable, if theoretically probable. Therefore, I claim the 1982 VRA is exogenous to U.S. steel firms' leverage because it was most likely implemented by the U.S. government.

The VRA went into effect on October 22, 1982. Its effect was comprehensive and immediate, limiting EC exports to 5.5% of the U.S. market beginning November 1, 1982 (October 22, 1982, *The New York Times*). In return, the U.S. firms dropped their unfair trade petitions and agreed to refrain from filing new cases until the VRA expired in January 1986. The VRA allowed U.S. firms to avoid further litigation costs and provided protection against all EC imports rather than a subgroup only (Moore, 1996). The large uncertainty perceived by the U.S. steel industry during the pre-VRA period was immediately resolved by the announcement of the 1982 VRA. In the absence of the VRA, the U.S. steel industry would have gone through cumbersome legal AD and CVD proceedings, whose effective outcomes are highly uncertain in the Knightian sense.

## 1.3.2 Empirical strategy and sample selection

### 1.3.2.1 Data

The data set used for the study consists of all nonfinancial firm-year observations in the annual Compustat database from 1971 to 2004. The period between 1978 and 1987 is for the 1982 VRA analysis and other periods are used for the placebo tests. In particular, the

period from 1978 to 1982 is designated as the pre-VRA period, and the period from 1983 to 1987 as the post-VRA period. I also require that all firm-year observations have valid data for book assets. All ratios are trimmed at the upper and lower one percentile to mitigate the effect of outliers and eradicate errors in the data. To incorporate the future expected profitability of a firm, I obtain median analyst forecasts of a firm's next-year earnings per share (EPS) from the IBES database and merge them with the accounting information obtained from the annual Compustat.

The definitions of each variable are detailed in Table 1.1 on page 40. The definitions are standard except for the time-varying forward-looking asset volatility. Following Faulkender and Petersen (2006, p. 60), I infer the time-varying asset volatility from the volatility of monthly equity returns. Departing from Faulkender and Petersen, I use the GARCH model to estimate the time-varying volatility of equity returns at a monthly frequency, because the GARCH-based asset volatility estimate is a more accurate forward-looking risk measure than the historical standard deviation approach. The GARCH model emphasizes a recent surprise close to the event for study, therefore it incorporates the realized asset volatility, as well as the forward-looking asset volatility when the 1982 VRA news arrives. The details of measuring the asset volatility using the GARCH model are in Appendix 1.E on page 36.

The distinction between the effect of risk and uncertainty is a critical study of this paper. A competing hypothesis against Knightian uncertainty is based on risk: the 1982 VRA reduces a U.S. steel firm's forward-looking asset risk, which completely explains away an increase in a steel firm's leverage in the post-VRA period. If the risk-based competing hypothesis is accurate, after controlling the forward-looking asset risk of the steel industry, the estimated effect of resolution of uncertainty in the post-VRA period should be statistically insignificant and/or its economic strength should be negligible. To rule out the risk-based competing hypothesis, the careful measurement of asset volatility is necessary.

### 1.3.2.2  Difference-in-Difference

The 1982 VRA is assumed to provide the resolution of uncertainty faced by the steel industry in the pre-VRA period, which is exogenous to the leverage decision of U.S steel producers. Precisely, the model used to estimate the causal effect of the resolution of uncertainty on U.S. steel firms' leverage is:

$$Y_{it} = \alpha_t + \alpha_i + \beta D_{it} + \Gamma' X_{it} + u_{it} \, , \tag{1.14}$$

where $(\alpha_t, \alpha_i, X_{it})$ are the firm-specific controls where $\alpha_t$ represents the time fixed effect, $\alpha_i$ the time-invariant firm fixed effect, and $X_{it}$ the time-varying observed firm characteristics.[14] The treatment group includes all U.S. steel producers. Three digit SIC codes are used to identify the steel manufacturers (331, 332) except firms producing non-ferrous metals, which are not covered under the 1982 VRA. For the control group, I select a subset of those firms similar to the steel manufacturers in the pretreatment firm characteristics. I use the standard propensity matching method for the selection of control firms.

The post-treatment indicator variable $D_{it}$ takes unity if a firm belongs to the U.S. steel industry in the post-VRA period (1983-1987). The difference-in-difference estimator, $\beta$, identifies the causal strength of $D_{it}$ on $Y_{it}$ while keeping $(\alpha_t, \alpha_i, X_{it})$ fixed. A vector of time-varying observed firm-specific controls $X_{it}$ includes firm size, profitability, tangibility, market-to-book ratio, time-varying asset volatility, annual stock return, and median analyst's forecast of profitability in year $t + 1$. $Y_{it}$ is the response variable: market leverage and book leverage. The error term $u_{it}$ is allowed to be serially correlated within firms and possibly heteroskedastic (Bertrand et al. 2004; Petersen 2009).

Selecting firm size, profitability, tangibility, and market-to-book ratio as firm-specific control variables is standard in the empirical capital structure literature (Rajan and Zingales 1995; Fama and French 2002; Leary and Roberts 2005). The rationale to include the other control variables is as follows. When good news arrives to the stock market, the steel industry's stock price tends to increase. Market leverage is defined as total debt divided by the sum of total debt and the market value of equity. Hence, an increase in a firm's stock price mechanically decreases its market leverage (Welch, 2004). When protected by the VRA, the U.S. steel industry is expected to be less competitive. Combining with the fact that the VRA provides the fixed term of protection, the future profitability of steel firms likely increases, and the future operating risk of the steel industry likely diminishes in the post-VRA period. To control these level effects, I include firm-specific annualized stock returns, median analyst forecast of next year's profitability, and time-varying forward-looking asset volatility as control variables. Then, the model (1.14) on the preceding page identifies the causal effect of resolution of uncertainty on leverage partialing out the other accompanying level effects described above.

---

[14] The linear model (1.14) identifies the causal strength of $D_{it}$ on $Y_{it}$ under the following assumptions:

1. Parallel-trends (conditional mean independence): $E[Y_{it}(0)|\alpha_t, \alpha_i, X_{it}, D_{it}] = E[Y_{it}(0)|\alpha_t, \alpha_i, X_{it}]$.

2. Time-invariant causal effect of $D_{it}$ on $Y_{it}$: $E[Y_{it}(1) - Y_{it}(0)|\alpha_t, \alpha_i, X_{it}] = \beta$.

3. Linearity: $E[Y_{it}(0)|\alpha_t, \alpha_i, X_{it}] = \alpha_i + \alpha_t + \Gamma' X_{it}$.

$Y_{it} = Y_{it}(0) + D_{it}(Y_{it}(1) - Y_{it}(0))$, where $\{Y_{it}(0), Y_{it}(1)\}$ are the standard potential outcomes notation for the response variable $Y_{it}$ (Rosenbaum and Rubin, 1983).

### 1.3.2.3   Constructing the matched sample

Table 1.2 on page 41 represents the comparison of the firm characteristic variables between the treatment and control group in the pre-VRA period (1978-1982). The Pre-match columns suggest that before matching, the means of observable firm characteristics in both groups are statistically significantly different, based on the sufficiently high $t$-statistics. Due to the imbalance of the pretreatment firm characteristics, a simple comparison of all non-steel firms to all firms in the steel industry is suspect.

For the correct causal inference of the treatment, we must compare treatment and control groups that are as similar as possible in pretreatment firm characteristics. If two groups have very different pretreatment firm characteristics, the prediction of counterfactual for the treatment group will be made using information of firms from the control group that looks very different from the treatment group (likewise for the prediction of counterfactual for the control group). Hence, the imbalance of firm characteristic variables between two groups leads to the estimation error of the causal effect of the treatment on the outcome.

The matched sample is constructed using the standard propensity score method (Rosenbaum and Rubin, 1983).[15] In the pre-VRA period, a logit model is estimated with the binary dependent variable, the indicator whether a firm belongs to the U.S. steel industry. Asset volatility, market-to-book ratio, annual stock return, and market leverage are the independent variables.[16] Then, I obtain the propensity scores of every firm in the sample from the estimated logit model. Finally, for each firm in the U.S. steel industry, I find up to four nearest non-steel firms in terms of the predicted propensity score with a caliper (i.e., propensity score distance) 0.05 to avoid the extreme counterfactuals.[17] In addition, I restrict the sample to firm-year observations that fall in the overlap between the domains of propensity scores for the treatment and control groups (i.e., common support condition). As a result, there are 41 unique steel-producing firms in the treatment group and 146 matched-control firms.

The post-match columns in Table 1.2 on page 41 present the results of the propensity score matching. After matching, based on the sufficiently low $t$-statistics, the firms in the treatment and control groups are statistically similar in terms of the means of each pretreatment firm-

---

[15]As a robustness check, I alternatively use Mahalanobis-metric direct matching with a caliper of 0.05 to construct the matched sample. The results are robust to this alternate matching method. The details are discussed in Section 1.5.2 and shown in Table 1.10.

[16]For a robustness check, as the independent variables, I use a different set of firm characteristics: size, profitability, tangibility, and market-to-book ratio. Table 1.10 shows that the results are robust to the use of alternative matching variables.

[17]I also use the different caliper sizes such as 1 and 0.08. The unreported results are robust to these alternative caliper sizes.

characteristic variable: tangibility, firm size, market-to-book ratio, profitability, time-varying volatility of asset, and simulated marginal corporate tax rate available from Graham (2000). The post-matching standard bias column in the table measures the difference in the means of each control variable divided by the standard deviation in the treatment group. Rubin and Thomas (2000) suggest that for reasonable balancing, the absolute standard bias of each covariate should be less than 0.25. Before matching, some of the absolute standard biases of control variables appear to be larger than 0.25. After matching, all absolute standard biases are less than 0.25. I perform the non-parametric test of the null hypothesis as to whether or not the distributions of the propensity scores of the treatment and control group are same after matching. The substantially high $p$-value (0.72) suggests to accept the null hypothesis. Finally, Figure 1.4 on page 42 illustrates the empirical distributions of each matching variable across the two groups after matching. As can be seen, those distributions are comparable. In sum, after matching, the pretreatment firm characteristics are balanced across the treatment and control groups, as desired.

## 1.4   Main Results

Table 1.5 on page 46 presents the main results of difference-in-difference (DID) analysis with the matched sample specified in (1.14) on page 17. The estimated coefficient of $D(i,t)$ (i.e., the DID estimate) in the table represents the estimated causal effect of resolution of uncertainty on steel firms' leverage relative to matched control firms. All standard errors are computed with clustering at the firm-level, thus robust to heteroscedasticity and within-firm serial correlation. In Models 1 and 2, all DID estimates are positive and statistically significant at the 5% level, observing sufficiently large $t$-statistics. That is, steel firms increase their leverage relative to matched control firms when uncertainty is resolved in the post-VRA period.

To gauge the economic strength of the causal effect of resolution of uncertainty on leverage, consider a median firm in the steel industry whose market leverage is 0.48 and book leverage is 0.26 in year 1982. The estimated coefficients of $D_{it}$ are 0.055 and 0.032, respectively. Scaling by its leverage, I find that when uncertainty is resolved, a median steel firm increases its market (book) leverage by $0.055/0.48 = 11.3\%$ ($0.032/0.26 = 12.3\%$) relative to a matched control firm. These results are economically strong.

By including the firm-specific control variables, level effects accompanied by the resolution of uncertainty as a result of the 1982 VRA are absorbed by those control variables. The signs of estimated coefficients of the control variables seem consistent with standard intuitions.

The size and tangibility of firms are positively related to leverage. Profitability enters with a negative sign. The estimated coefficient of median analyst forecast of earnings per share as a proxy for the next year's profitability is negative and statistically significant at the 5% level. The negative sign on profitability seems counterintuitive to the standard trade-off theory. However, the empirical studies typically find a negative relation between profitability and leverage ([see Frank and Goyal (2008) and the references therein]). Market-to-book ratio is negatively related to market leverage. It is consistent with trade-off theories because growth firms lose more of their value when they go into distress. By definition, market leverage is inversely related to an increase in the market value of equity (Welch, 2004). Hence, it seems reasonable that a firm's stock return is also negatively related to its market leverage. Forward-looking asset volatility is also negatively related to leverage. It is also consistent with trade-off theories because firms with more volatile assets will have higher probabilities and expected costs of bankruptcy.

Models 3 and 4 in Table 1.5 on page 46 present the results when the excess price-cost margin of a firm (henceforth, EPCM) is included as an additional control variable for the firm-level market power. Following Gaspar and Massa (2006), the EPCM of a firm is computed as the difference between a firm's price-cost margin (Lerner index) and the average operating profit margin of its industry. The price-cost margin is defined as sales minus cost of goods sold minus selling, general, and administrative expenses divided by total asset. The estimated coefficient of EPCM is positive, although statistically insignificant at the 10% level. When the steel industry is protected by import quotas, these firms face less competition and the price-cost margin of a steel firm increases accordingly. The steel firms facing a less risky environment tend to take on more leverage. Therefore, there is likely a positive relation between the EPCM of a firm and its leverage.

In addition to leverage, I estimate the causal effect of resolution of uncertainty on a steel firm's cash holding ratio (cash/asset). I propose that cash is not only a less risky asset, but also a less uncertain asset than debt. Therefore, the prediction is that when uncertainty is resolved, a steel firm likely reduces its cash holding to a matched control firm. Table 1.7 on page 48 presents the results that verify the prediction. The estimated coefficient of $D(i,t)$ (the DID estimate) is negative (-0.22) and statistically significant at the 10% level. Consider a median firm whose cash holding ratio is 0.06 in the pre-VRA period. When uncertainty is resolved, a median firm reduces its cash holding ratio by $0.022/0.06 = 36.7\%$ compared to a matched control firm. This effect is also economically strong.

One of the key assumptions to ensure the internal validity of the difference-in-difference methodology is the parallel-trends assumption (Roberts and Whited, 2011). In the absence

of the 1982 VRA, the average change in leverage would have been the same for both the treatment and control groups. Because we never observe the true error terms $u_{it}$ of the model (1.14) on page 17, we cannot perform the formal statistical test to ensure the necessary and sufficient conditions for the parallel-trends assumption. But it is still possible to check the necessary condition by comparing trends in the outcome variable across the treatment and control groups in the pre-treatment period (Roberts and Whited, 2011). Since I have multiple pretreatment firm-year observations (1979-1982), I can perform a two-sample test to compare the average growth rates (average slopes) of the treatment and matched control groups.

Table 1.4 on page 44 presents the results of these tests. For example, in the second row the difference in the means of market-leverage growth rate of the two groups in 1979-1980 is 0.044. The high $p$-value from the two-sample $t$-test suggests that the the difference is statistically insignificant. In addition, the sufficiently high $p$-value from rank-sum test also suggests the distributions of market-leverage growth rate of the treatment and control group are statistically indistinguishable. I perform similar tests for book leverage and cash holding ratio in the pretreatment periods (1978-1982). All tests suggest that the average trends of leverage of the treatment and control group in the pre-VRA period are statistically similar, based on sufficiently high $p$-values of both $t$-tests and rank-sum tests. Statistically, the necessary condition for the parallel-trends assumption appears to be satisfied, as desired.

Figure 1.5 on page 45 depicts the average market-leverage trends of U.S. steel firms and the matched control firms. In the pre-VRA period (1979-1982), the trends across the two groups are comparable. It is expected from the results of the formal statistical tests presented in Table 1.4. In the post-VRA period (1983-1987), the two lines diverges noticeably. While the mean market leverage of the matched control group remains the same, that of the treatment group sharply increases. This difference in the average market leverage between the treatment and matched control provides the (informal) time-varying estimates of the causal effect of VRA on the market leverage of the U.S. steel industry in the post-VRA period. The discrepancy between the two lines in the post-VRA period in Figure 1.5 suggests a large positive effect of the VRA (i.e., the resolution of uncertainty) on the leverage of steel firms relative to the matched control firms. The DID estimates (0.055) for market leverage presented in Table 1.5 on page 46 is conceptually equivalent to the time-averaged differences between the two lines. Note that to generate this figure, I require that all firms in the two groups have full firm-year observations in the pre- and post-VRA period (10 years).

## 1.5 Robustness Checks

### 1.5.1 Collapsing the data into two periods

Bertrand et al. (2004) suggest that the most conservative way to deal with the within-firm serial correlation of errors is to ignore the time-series information. Following Bertrand et al., I time-average the data before and after the 1982 VRA and run analysis (1.14) on this averaged outcome variable in a panel of length 2 (Bertrand et al., 2004, Section 4.5.). Table 1.8 on page 49 reports the results of this exercise. Compared to the main results in Table 1.5 on page 46, in Model 1 for market leverage, the DID estimate (i.e., the estimated coefficient of $D(i)$) increases from 0.05 to 0.057, while its statistical significance remains similar. In Model 2 for book leverage, the DID estimate decreases from 0.032 to 0.029, and its statistical significance decreases from 0.03 to 0.12 in terms of $p$-value. Models 3 and 4 present the results of similar analyses when excess price-cost margin is employed as an additional control variable. The results of Models 3 and 4 are similar to those of Models 1 and 2.

Losing the statistical significance of the DID estimates in this exercise is expected because collapsing the data into two periods substantially reduces the sample size, while the number of covariates remains the same. As such, I lose a degree of freedom, which leads to the weaker statistical significance of the DID estimates. This aspect is also pointed out in Bertrand et al. (2004).[18] However, the signs of DID estimates are consistent with the main results presented in Table 1.5, and the magnitude of them is also comparable. The results suggest that the causal effect of the resolution of uncertainty provided by the 1982 VRA on a steel firm's leverage remains economically strong in the most conservative setup.

### 1.5.2 Placebo tests and alternative matching methods

Could the observed increase in a steel firm's leverage responding to the resolution of uncertainty simply be random? A way to answer to this question is to generate "placebo-shocks" as in Bertrand et al. (2004). Specifically, I conduct three placebo tests in which I pretend the VRA happens, instead of in 1982, at different times: 1979, 1994, and 2002. The results appear in Table 1.9 on page 50. In the 1979 and 1994 placebo-VRAs, none of the DID estimates is statistically or economically significant, observing the weak economic strength and the sufficiently small $t$-statistics. In the 2002 placebo-VRA, all of the DID estimates are

---

[18]Bertrand et al. (2004) states that "$\cdots$ The downside of these procedures (both raw and residual aggregation) is that their power is quite low and diminishes fast with sample size."

statistically insignificant at the 10% level. Although their economic strength seems nonnegligible, they enter with a negative sign. As a result, none of the placebo VRAs can generate the economically and statistically effect as significant as that estimated using the actual 1982 VRA shock.

Next, I show that the main results are robust to alternative matching methods. To begin, I match on the same pretreatment characteristics as shown in Section 1.3.2.3: the volatility of asset, stock return, market-to-book ratio, and market leverage ratio. Here, I use a Mahalanobis-metric to measure the distance of two firms with the chosen characteristics. Each steel firm is matched to the four nearest non-steel firms within a caliper of 0.05.[19] The alternative matching (I) columns in Table 1.10 on page 51 present the results of difference-in-difference analysis with the matched sample using this matching method. The DID estimates are similar to those presented in Table 1.5 in terms of statistical and economic significance.

Finally, to check the sensitivity of selection of matching variables, I use a different set of characteristic variables to estimate the propensity scores: tangibility, size, profitability, and market-to-book ratio. As before, each steel firm is matched up to four nearest non-steel firms in terms of the predicted propensity score with a caliper 0.05. The alternative matching (II) columns in Table 1.10 present the results. All of the DID estimates are economically strong and statistically significant at the 10% level. The economic magnitude and statistical significance of DID estimates are similar to those of the main results presented in Table 1.5.

## 1.6    Conclusion

This paper shows that the Knightian uncertainty perceived by a manager and her aversion toward it are important drivers of leverage. The model incorporates an ambiguity-averse manager and predicts a substantially lower optimal leverage ratio than its traditional counterpart. Quantitatively, the effect of uncertainty on leverage is higher than that of risk. Moreover, the firm's leverage strictly decreases with uncertainty, whereas it can be non-monotone in risk even when the manager is risk-averse. Therefore, I argue that uncertainty provides a more plausible explanation for firms taking on low leverage than risk.

I find a causal relationship between uncertainty and leverage. Prior to the 1982 VRA, U.S.

---

[19]This caliper is chosen following Rubin and Thomas (2000). They advise that in the case of Mahalanobis-distance matching, when the variance of linear propensity score in the treatment group is less than twice as large as that in the control group, a caliper of 20% standard deviation of propensity score of the treatment group removes 98% of the bias in a normally distributed covariates. I find that the standard deviation of linearized propensity scores of treatment group is 0.27, whereas that of control group is 0.28. Hence, it seems proper to use $0.2 \times 0.27 \approx 0.05$ as a caliper.

steel firms likely perceived high Knightian uncertainty about the likelihood of outcomes of AD and CVD legal proceedings. The source of uncertainty was that the ITA had considerable discretion and often made highly arbitrary injury determinations. I use the 1982 VRA as an exogenous reduction in Knightian uncertainty faced by firms in the U.S. steel industry. Using a difference-in-difference methodology with the matched sample, I find that when uncertainty is resolved, a median firm in the steel industry increases its market leverage by 11.5% and its book leverage by 12.3% as compared to a matched control firm. These results are economically strong and statistically significant. Most importantly, the results are not explained away by changes in traditional factors such as risk and future profitability.

The 1982 VRA was implemented by the U.S. government because of its own strong political incentives. Therefore, the internal validity of the difference-in-difference methodology is not threatened by steel firms lobbying for the VRA. I also show that the results are robust to placebo shocks and to alternative matching methods.

This paper is, to the best of my knowledge, the first paper that investigates the effect of Knightian uncertainty on optimal capital structure. It contributes to the capital structure literature by providing a formal model and empirical evidence showing that Knightian uncertainty provides a sensible explanation for firms taking on low leverage.

# 1.A  Ellsberg Paradox and Reasonable Ambiguity aversion Parameter

Agents are given the opportunity to bet on the draw of a ball from one of two urns. Urn A has 50 red and 50 black balls, while urn B has 100 balls in an unknown mix of red and black. First, the agents are offered a choice between two bets: (a) \$100 if the ball drawn from urn A is red and nothing if it is black, or (b) \$100 if the ball drawn from urn B is red and nothing if it is black. In the first experiment, the agents generally prefer the first bet over the second. Therefore, the predicted probability of red in urn B must be strictly less than 0.5. Next, the agents are offered a choice between two new bets: (a) \$100 if the ball drawn from urn A is black and nothing if it is red, (b) \$100 if the ball drawn from urn B is black and nothing if it is red. Again the first bet has been generally preferred in experiments. It implies that the predicted probability of black in urn B must be less than 0.5, i.e., the predicted probability of red in urn B must be strictly higher than 0.5. Indeed, this probability assessment is inconsistent since one cannot simultaneously assign to the event "red from urn B" a probability that is strictly less and also more than 0.5. Ellsberg (1961) interprets this behavior as an aversion to the ambiguity about the odds for urn B. Therefore the agents prefer to bet on events with known odds than on events with ambiguous odds. This phenomenon is referred to as ambiguity aversion (Ellsberg, 1961).

Consider the above Ellsberg experiment. Assume that the risk-averse agent display the constant absolute risk aversion (CARA) $u(x) = 1 - e^{-\rho x}$ and the ambiguity averse agent display the constant absolute ambiguity aversion (CAAA) $\phi(x) = 1 - e^{-\alpha x}$. The agent knows that the distribution of red and black balls in Urn A is (1/2,1/2). When she evaluates a "risky" bet on Urn A, her certainty equivalent is equal to, given the CARA coefficient $\rho = 2$,

$$u^{-1}\left(\frac{1}{2}u(100) + \frac{1}{2}u(0)\right) = 0.28. \tag{1.15}$$

The agent believes that there are two possible equally likely distributions: her set of priors $\Pi = \{(0,1),(1,0)\}$ and $\mu = (1/2,1/2)$. But she is averse to this uncertainty and displays CAAA. When she evaluates a "ambiguous" bet on Urn B, her non-ambiguity equivalent – the payoff amount she would accept in lieu of the ambiguous situation – can be computed as, given the CAAA coefficient $\alpha = 4$,

$$\phi^{-1}\left(\frac{1}{2}\phi(1) + \frac{1}{2}\phi(0)\right) = 0.17. \tag{1.16}$$

The difference between the certainty equivalent (1.15) and the non-ambiguity equivalent (1.16) is a measure of ambiguity premium (Klibanoff et al., 2005; Ju and Miao, 2007). Camerer (1999) reported that the ambiguity premium is typically on the order of 10 - 20% of the expected value of a bet in the behavioral setting (in this example, $0.2 \times 0.5 = 0.1$). Therefore, the choice of CAAA ambiguity aversion parameter seems reasonable with respect to the CARA risk aversion parameter.

## 1.B   Proofs

### 1.B.1   Proof of Lemma 1.1

*Proof.* Assume $\Pi$ is compact so that I can interchange differentiation and integration.[20]

$$\mathbb{E}^{\mu}\left[\phi^{'}(V(B|F))V_{B}(B|F)\right] = \mathbb{E}^{\mu}(\phi^{'}(V(B|F))) \cdot \mathbb{E}^{\mu}\left[\frac{\phi^{'}(V(B|F))}{\mathbb{E}^{\mu}(\phi^{'}(V(B|F)))}\frac{\partial V(B|F)}{\partial B}\right].$$

Since,

$$\mathbb{E}^{\mu}\left[\frac{\phi^{'}(V(B|F))}{\mathbb{E}^{\mu}\phi^{'}(V(B|F))}\right] = 1.$$

I can define $Z = \frac{\phi^{'}(V(B|F))}{\mathbb{E}^{\mu}\phi^{'}(V(B|F))}$ as the density (Radon-Nykodym derivative) of $\mu^{*}$ with respect to $\mu$. That is,

$$d\mu^{*} = \frac{\phi^{'}(V(B|F))}{\mathbb{E}^{\mu}(\phi^{'}(V(B|F)))} \cdot d\mu$$

Thus, the first order condition can be rewritten as:

$$\mathbb{E}^{\mu^{*}}[V_{B}(B|F))] = \int_{\Pi} V_{B}(B|F)d\mu^{*}(F) = 0.$$

$\square$

### 1.B.2   Proof of Corollary 1.1

*Proof.* Assume a finite dimensional parameter space $\Theta$. Any parametric setting, $\Pi := \{F(\cdot|\theta)\}_{\theta \in \Theta}$, therefore $\Pi$ is fully indexed by $\theta$. Assume also $\Pi$ is compact.

---

[20]In a case where $\Pi$ is not compact, I use the Lebesgue dominated convergence theorem. Provided there exists a integrable function $g$ such that $|\Phi'(\cdot)| < g(\cdot)$, I can interchange integration and differentiation (Folland, 1999).

The first-order condition is:

$$\int_{\Pi} \phi'(V(B|\theta))\frac{\partial V(B|\theta)}{\partial B}d\mu(\theta) = \int_{\Pi} \phi'(V(B|\theta))\left(kBf(B|\theta) - \tau\int_{B}^{\infty} f(x)dx\right)d\mu(\theta)$$

$$= kB\int_{\Pi} f(B|\theta)\phi'(V(B|\theta))d\mu(\theta) - \tau\int_{\Pi} \phi'(V(B|\theta))\left(\int_{[B,\infty]} f(x|\theta)dx\right)d\mu(\theta)$$

$$= kB\int_{\Pi} f(B|\theta)\phi'(V(B|\theta))d\mu(\theta) - \tau\int_{\Pi} \phi'(V(B|\theta))\left(1 - F(B|\theta)\right)d\mu(\theta)$$

$$= kB\int_{\Pi} f(B|\theta)\frac{\phi'(V(B,\theta))}{\int_{\Pi} \phi'(V(B,\theta))d\mu(\theta)}d\mu(\theta) - \tau\left(1 - \int_{\Pi} F(B|\theta)\frac{\phi'(V(B|\theta))}{\int_{\Pi} \phi'(V(B,\theta))d\mu(\theta)}d\mu(\theta)\right)$$

$$= kB\int_{\Pi} f(B|\theta)d\mu^*(\theta) - \tau\left(1 - \int_{\Pi} F(B|\theta)d\mu^*(\theta)\right)$$

$$= kB\,\mathbb{E}^{\mu^*}[f(B)] - \tau\left(1 - \mathbb{E}^{\mu^*}[F(B)]\right) = 0\,.$$

That is, the manager balances the trade-off between the ambiguity-adjusted marginal tax benefit of debt and ambiguity-adjusted marginal default cost of debt.

In addition, the sufficient condition for the optimality holds. To see this, obtain the second-order condition:

$$\frac{\partial}{\partial B}\left(\int_{\Pi} \phi'(V(B|\theta))\frac{\partial V(B|\theta)}{\partial B}d\mu(\theta)\right)$$

$$= \int_{\Pi} \left(\phi''(V(B|\theta))\frac{\partial V(B|\theta)}{\partial B} + \phi'(V(B|\theta))\frac{\partial^2 V(B|\theta)}{\partial B^2}\right)d\mu(\theta)$$

$$= \int_{\Pi} -\alpha\phi'(V(B|\theta))\frac{\partial V(B|\theta)}{\partial B}d\mu(\theta) + \int_{\Pi} \phi'(V(B|\theta))\frac{\partial^2 V(B|\theta)}{\partial B^2}d\mu(\theta)$$

$$= \int_{\Pi} \phi'(V(B|\theta))(-kf(B|\theta) - kBf'(B|\theta) - \tau f(B|\theta))d\mu(\theta) < 0\,.$$

In the second line, I use the property of CAAA such that $\phi''(x) = -\alpha\phi'(x)$, then the first term vanishes due to the FOC. The last line evidently holds because $\phi'(V) > 0$ and $B < \theta$. $\qquad\square$

## 1.B.3   Proof of Proposition 1.1

*Proof.* Let $G(\alpha, B, \Pi) = \int_{\Pi} \phi'(V(B|\theta))\frac{\partial V(B|\theta)}{\partial B}d\mu(\theta) = 0$. Since the sufficient condition for the optimality of $B$ holds (*i.e.*, $\frac{\partial G}{\partial B} < 0$), by the implicit function theorem, it is enough to

show $\frac{\partial G}{\partial \alpha} < 0$. Because $\phi'(x) = \alpha e^{-\alpha x}$ and $\frac{\partial}{\partial \alpha}\phi'(x) = \frac{1}{\alpha}\phi'(x) - \phi'(x)x$,

$$
\begin{aligned}
\frac{\partial G}{\partial \alpha} &= \int_\Pi \left(\frac{1}{\alpha}\phi'(V(B|\theta)) - \phi'(V(B|\theta))V(B|\theta)\right)(-kBf(B) + \tau(1 - F(B)))d\mu(\theta) \\
&= -\int_\Pi \phi'(V(B|\theta))V(B|\theta)(-kBf(B) + \tau(1 - F(B)))d\mu(\theta) < 0\,.
\end{aligned}
$$

In the first line, I use the FOC (1.8) on page 8. The first term of the second line also vanishes because of FOC. Observe $V(B, \theta) > 0$ and $V'(B, \theta) > 0$ for all $\theta$ in $\Pi$. Therefore, the last line is valid.

Next, let the reference parameter be $\theta_0$ and the amount of uncertainty be $\delta$. Accordingly, the set of priors is defined as $\Pi = [\overline{\theta}, \underline{\theta}]$, where $\overline{\theta} = \theta_0 + \delta$ and $\underline{\theta} = \theta_0 - \delta$. Assume the manager puts the uniform belief $\mu(d\theta) = \frac{1}{2\delta}d\theta$ on $\Pi$. Then, at the optimal $B$ satisfying (1.8),

$$
\begin{aligned}
\frac{\partial}{\partial \delta}&\left(\int_{\theta_0-\delta}^{\theta_0+\delta}(-kBf(B|\theta) + \tau\overline{F}(B|\theta))\phi'(V(B|\theta))\frac{1}{2\delta}d\theta\right) \\
&= (-kBf(B|\overline{\theta}) + \tau\overline{F}(B|\overline{\theta}))\phi'(V(B|\overline{\theta})) + (-kBf(B|\underline{\theta}) + \tau\overline{F}(B|\underline{\theta}))\phi'(V(B|\underline{\theta})) \\
&\quad -\frac{1}{\delta}\int_\Pi(-kBf(B|\theta) + \tau\overline{F}(B|\theta))\phi'(V(B|\theta))\frac{1}{2\delta}d\theta < 0\,.
\end{aligned}
$$

where $\overline{F}(B|\theta) = 1 - F(B|\theta)$. At the optimal $B$ that satisfies (1.8), the last term in the second line vanishes due to the FOC, and $-kBf(B|\overline{\theta}) + \tau\overline{F}(B|\overline{\theta}) > 0$ and $-kBf(B|\underline{\theta}) + \tau\overline{F}(B|\underline{\theta}) < 0$. Finally, observe $\phi'(v)$ is (exponentially) monotone decreasing function and $V(B|\underline{\theta}) < V(B|\overline{\theta})$, where $\underline{\theta} < \overline{\theta}$. Therefore, the last line is valid. $\qquad\square$

### 1.B.4  Proof of Proposition 1.2

**Lemma 1.2.** *When a (risk-averse) manager faces uncertainty, the ambiguity-neutral manager combines the uncertainty about the distribution and risk of the distribution into the variance of a new density. Precisely, let $\Pi = \{F(\cdot|\theta)\}_{\theta \in \Theta}$ be the compact set of prior and $\mu$ associated belief over $\Pi$. Then, there exists a density $\widehat{f}$ given by:*

$$
\widehat{f}(x) = \int_\Pi f(x|\theta)d\mu(\theta)\,,
$$

*such that*

$$
\int_\Pi\left(\int_{\underline{x}}^{\overline{x}} u(x)f(x|\theta)dx\right)d\mu(\theta) = \int_{\underline{x}}^{\overline{x}} u(x)\widehat{f}(x)dx\,.
$$

*That is, an ambiguity-neutral manager acts behaviorally the same as a Bayesian manager who puts the prior $\mu(d\theta)$ on the parameter of $f(x|\theta)$.*

*Proof.* By Fubini's theorem,

$$\int_\Pi \left( \int_{\underline{x}}^{\overline{x}} u(x) f(x|\theta) dx \right) d\mu(\theta) = \int_{\underline{x}}^{\overline{x}} u(x) \left( \int_\Pi f(x|\theta) d\mu(\theta) \right) dx = \int_{\underline{x}}^{\overline{x}} u(x) \widehat{f}(x) dx$$

$\square$

*Proof of Proposition 1.2 on page 11.* According to Lemma 1.2 on the preceding page, in the presence of uncertainty, a risk-averse but ambiguity-neutral manger constructs the new compounding distribution $\widehat{f}(x)$ combining the uncertainty and risk of the distribution. For the simplicity of notation, use $f(x)$ instead of $\widehat{f}(x)$.

The risk-averse but ambiguity-neutral manager's optimality condition for $B$ is:

$$f(B)(u(B) - u(B(1-k))) = \tau \int_B^\infty u'((1-\tau)x + \tau B) f(x) dx, \tag{1.17}$$

where $u(\cdot)$ CARA utility function such that $u'(x) > 0$, $u''(x) < 0$ and $\alpha = -u''/u'$. Because $u(x)$ is continuous and differentiable on $[(1-k)B, B]$, by the mean value theorem there exists $B_0 \in [(1-k)B, B]$ such that:

$$kB u'(B_0) f(B) = \int_B^\infty u'((1-\tau)x + \tau B) f(x) dx. \tag{1.18}$$

Let
$$G(B, \sigma) = -kB u'(B_0) f(B) + \int_B^\infty u'((1-\tau)x + \tau B) f(x) dx = 0. \tag{1.19}$$

Then, by the implicit function theorem,

$$\frac{\partial B}{\partial \sigma} = -\frac{\frac{\partial G}{\partial \sigma}}{\frac{\partial G}{\partial B}}.$$

From (1.19), we have:

$$-\frac{\partial G}{\partial B} = k u'(B_0) f(B) + k B u'(B_0) \frac{\partial f(B)}{\partial B} + \tau u'(B) - \tau^2 \int_B^\infty u''((1-\tau)x + \tau B) f(x) dx > 0$$

and
$$\frac{\partial G}{\partial \sigma} = -k B u'(B_0) \frac{\partial f(B)}{\partial \sigma} + \tau \frac{\partial}{\partial \sigma} \int_B^\infty u'((1-\tau)x + \tau B) f(x) dx \tag{1.20}$$

Since $u'(t) > 0$ for all $t \in [B, \infty]$, we have:

$$\frac{\partial G}{\partial \sigma} < -k \, B \, u'(B_0) \frac{\partial f(B)}{\partial \sigma} + \sup_{t \in [B, \infty)} u'((1 - \tau)x + \tau B) \cdot \frac{\partial}{\partial \sigma} \int_B^\infty f(x) dx \,.$$

Assuming $B < \mathbb{E}X = \theta$, $\frac{\partial}{\partial \sigma} \left( \int_B^\infty f(x) dx \right) < 0$ for all $0 < \sigma < +\infty$. Hence, (i) if:

$$\frac{\partial f(B)}{\partial \sigma} > 0 \iff \sigma < \theta - B \,,$$

then combining with $-\frac{\partial G}{\partial B} > 0$:

$$\frac{\partial G}{\partial \sigma} < 0 \iff \frac{\partial B}{\partial \sigma} < 0 \,.$$

Otherwise, let a solution of equation (1.20) be $\bar{\sigma}$. Then, if $\theta - B < \sigma < \bar{\sigma}$,

$$\frac{\partial G}{\partial \sigma} < 0 \iff \frac{\partial B}{\partial \sigma} < 0 \,,$$

but (ii) if $\sigma \geq \bar{\sigma}$, then:

$$\frac{\partial G}{\partial \sigma} \geq 0 \iff \frac{\partial B}{\partial \sigma} \geq 0 \,.$$

$\square$

## 1.B.5 Derivation of the condition (1.11)

Assume $\Pi$ is compact. Let the optimal leverage taken by the risk-averse but ambiguity-neutral manager be $B_R$, which then satisfies the first-order condition:

$$G_R(B_R) = -K B_R u'(B_0) \int_\Pi f(B_R|\theta) d\mu(\theta) + \tau \int_\Pi \int_{B_R}^{\bar{x}} u'((1 - \tau)x + \tau B_R) f(x|\theta) dx d\mu(\theta) = 0 \,, \tag{1.21}$$

where $B_0 \in [(1 - k)B_R, B_R]$ defined in (1.18). Let the optimal leverage taken by the risk-neutral but ambiguity-averse manager be $B_A$, which then satisfies the first-order condition:

$$G_A(B_A) = -kB \int_\Pi \phi'(V(B_A|\theta)) f(B_A|\theta) d\mu(\theta) + \tau \int_\Pi \phi(V(B_A|\theta)) \int_{B_R}^{\bar{x}} f(x|\theta) dx d\mu(\theta) = 0 \,. \tag{1.22}$$

Assuming $B < \theta$, $\frac{\partial G_R}{\partial B} < 0$ and $\frac{\partial G_A}{\partial B} < 0$. That is, $G_R(B)$ and $G_A(B)$ are monotonically decreasing in $B$.

Since $G_A(B)$ is a monotone decreasing function on $B$, it is sufficient to show that

$G_A(B_R) < 0$ in order to prove $B_A < B_R$.

$$G_A(B_R) = -kB_R \int_\Pi \phi'(V(B_R|\theta))f(B_R|\theta)d\mu(\theta) + \tau \int_\Pi \phi'(V(B_R|\theta)) \left( \int_{B_R}^\infty f(x|\theta)dx \right) d\mu(\theta).$$

By the mean value theorem, there exists a $\theta_1 \in \Pi$ such that:

$$
\begin{aligned}
G_A(B_R) &= -kB_R\phi'(V(B_R|\theta_1)) \int_\Pi f(B_R|\theta)d\mu(\theta) \\
&\quad + \tau \int_\Pi \phi'(V(B_R|\theta)) \left( \int_{B_R}^\infty f(x|\theta)dx \right) d\mu(\theta).
\end{aligned}
\tag{1.23}
$$

According to (1.21),

$$kB_R \int_\Pi f(B_R|\theta)d\mu(\theta) = \tau \int_{B_R}^\infty \frac{u'((1-\tau)x + \tau B_R)}{u'(B_0)} \left( \int_\Pi f(x|\theta)d\mu(\theta) \right) dx.$$

Plug it into (1.23),

$$
\begin{aligned}
G_A(B_R) &= -\tau\phi'(V(B_R|\theta_1)) \int_{B_R}^\infty \frac{u'((1-\tau)x + \tau B_R)}{u'(B_0)} \left( \int_\Pi f(x|\theta)d\mu(\theta) \right) dx \\
&\quad + \tau \int_\Pi \phi'(V(B_R|\theta)) \left( \int_{B_R}^\infty f(x|\theta)dx \right) d\mu(\theta) \\
&= \left( -\phi'(V(B_R|\theta_1)) \frac{u'((1-\tau)x_0 + \tau B_R)}{u'(B_0)} + \phi'(V(B_R|\theta_2)) \right) \\
&\quad \times \int_\Pi \left( \int_{B_R}^\infty f(x|\theta)dx \right) d\mu(\theta),
\end{aligned}
\tag{1.24}
$$

where $\theta_2 \in \Pi$ and $x_0 \in [B_R, \bar{x}]$. I use the mean value theorem twice in the the last line.

According to (1.24), for any given $B_R$, $B_A < B_R$ if and only if:

$$\frac{u'((1-\tau)x_0 + \tau B_R)}{u'(B_0)} > \frac{\phi'(V(B_R|\theta_2))}{\phi'(V(B_R|\theta_1))}.
\tag{1.25}$$

Assuming $\phi(\cdot)$ is the CAAA and $u(\cdot)$ is the CARA preference, (1.25) is equivalent to:

$$\rho((1-\tau)x_0 + \tau B_R - B_0) < \alpha(V(B_R|\theta_2) - V(B_R|\theta_1)),
\tag{1.26}$$

where $\alpha$ is the coefficient of absolute ambiguity aversion, and $\rho$ is the coefficient of absolute risk aversion.

Note that the condition (1.26) is more easily satisfied when keeping all the other things

fixed, (i) the ambiguity aversion ($\alpha$) is higher than risk aversion ($\rho$), (ii) the dispersion of $\Pi$ (*i.e.*, $\|\theta_1 - \theta_2\| = \delta$) is higher relative to risk ($\sigma$), because a decrease in $\sigma$ decreases $x_0 \in [B_R, \bar{x}]$ and increases $B_0 \in [(1 - k)B_R, B_R]$, which the left-hand side of the inequality decreases.

## 1.C   Ambiguity-averse Manager: Numerical Examples

To compute the optimal debt of the ambiguity-augmented model, I make the following assumptions:

1. The Knightian manager is risk-neutral and knows $X \sim N(\theta, \sigma^2)$. She is certain about the variance $\sigma^2$, but uncertain about the mean $\theta$. Instead, she collects all candidates of distribution $F_i(x)$ for $X$ and constructs $\Pi = \{F_i(x)\}_{i=1}^n$.

2. Suppose all distributions $\{F_i(x)\}_{i=1}^n$ are Gaussian distributions with a known and common standard deviation $\sigma$. Then $F_i(x) = F(x; \theta_i, \sigma)$ is fully indexed by $\theta_i$ when $\sigma$ is known. The set $\Pi$ is assumed to be finite, so I can write $\Pi = \{\theta_i\}_{i=1}^N$ without loss of any information.

3. The manager's attitude toward ambiguity is characterized by the constant absolute ambiguity aversion (CAAA) function as (1.9) on page 9 in Definition 1.

4. The dispersion of $\Pi$, denoted as $\|\Pi\|$, represents the amount of uncertainty the manager faces. Under assumption 2,

$$\Pi = \{\theta_{min}, \cdots, \theta_{max}\}, \qquad \theta_{min} < \theta_{max}.$$

   Then, the length of the interval $|\theta_{max} - \theta_{min}|$ can be a representation of $\|\Pi\|$. A larger $\|\Pi\|$ implies that the manager believes that the true mean belongs to a larger interval (*i.e.*, she confronts higher uncertainty). Fixing the average of $\{\theta_i\}_{i=1}^n$, as $\|\Pi\|$ increases, the lowest possible mean $\theta_{min}$ decreases. Although $\theta_{max}$ also increases as $\|\Pi\|$ increases, the manager's ambiguity aversion forces her to pay higher attention to the worst case ($\theta_{\min}$) than the best ($\theta_{max}$).

5. $N$, the number of possible prior, is assumed to be 3 and the firm's volatility $\sigma$ is 0.3. The manager believes all three priors in $\|\Pi\|$ are equally likely. That is,

$$\Pi = \{\theta_1, \theta_2, \theta_3\}, \quad \sigma = 0.3, \quad \text{and} \quad \mu(\theta_i) = \frac{1}{3}, \quad \text{where } i = 1, 2, 3 \qquad (1.27)$$

33

such that $\|\Pi\| = |\theta_1 - \theta_3|$.

6. The optimal leverage ratio is defined as:

$$\text{Levarage ratio} = \frac{B^*}{E(B^*) + B^*}, \qquad (1.28)$$

where $B^*$ is the optimal amount of debt computed from a relevant first-order condition, and $E(B^*)$ is market value of equity at $B^*$ using (1.3) on page 6.

Under these assumptions, the ambiguity-averse manager chooses the optimal debt level to satisfy:

$$\max_{B>0} \sum_{i=1}^{3} \mu(\theta_i) \cdot \phi(V(B, \theta_i)) \qquad (1.29)$$

with respect to (1.9) on page 9, (1.27) and

$$V(B, \theta_i) = \left(\frac{1}{1+r}\right)\left(\int_0^B x(1-k)\,f(x; \theta_i, \sigma)dx + \int_B^\infty \left((1-\tau)x + \tau B)\right)f(x; \theta_i, \sigma)dx\right),$$

where $f(x; \theta_i, \sigma)$ is the normal density with the mean $\theta_i \in \Pi$ and a known standard deviation $\sigma$.

According to Proposition 1.1 on page 7, the first-order condition of (1.29) is:

$$\sum_{i=1}^{3} \mu^*(\theta_i)\frac{\partial V(B, \theta_i)}{\partial B} = 0, \qquad (1.30)$$

where the ambiguity-adjusted belief is:

$$\mu^*(\theta_i) \equiv \frac{\mu(\theta_i)\phi'(V(B, \theta_i))}{\sum_j^3 \mu(\theta_j)\phi'(V(B, \theta_j))}.$$

Then the optimal leverage $B$ satisfies:

$$kBf^*(B) = \tau(1 - F^*(B)),$$

where $f^* = \sum_i^3 \mu^*(\theta_i)f(x|\theta_i)$ is a mixture of normal densities $f(x|\theta_i)$ with ambiguity-adjusted belief $\mu^*(\theta_i)$. It is a discrete counterpart of Corollary 1.1 on page 8.

For the numerical computations, I set the marginal corporate tax rate as $\tau = 0.35$ and the proportional deadweight cost of default as $k = 0.4$. The particular choice of tax rates are applicable to the recent period 1990 - 2009 after the 1986 Tax Reform Act (Graham,

2000). I also fix the set of priors and associated belief as:

$$\Pi = \{0.7, 1, 1, 3\} \qquad \mu(\theta_i) = \frac{1}{3} \qquad i = 1, 2, 3\,, \tag{1.31}$$

such that the amount of uncertainty is fixed as $\|\Pi\| = 0.3$. The other parameters are set as before: $\sigma = 0.3$, $\tau = 0.35$, and $k = 0.4$. I now change only the manager's degree of ambiguity aversion $(\alpha)$ and numerically compute the optimal leverage decisions of the ambiguity-aversion augmented model with varying $\alpha$. Figure 1.1 on page 38 illustrates the results. As expected, the result is consistent with Proposition 1.1 on page 9. As the manager's ambiguity aversion increases ($\alpha$ rises), she takes on less leverage. In Figure 1.1, the ambiguity-averse manager ($\alpha = 8$) takes substantially less leverage (0.64) than the ambiguity-neutral manager (0.80).

I turn to compute the optimal leverage ratio with respect to the amount of uncertainty perceived by the manager. I specify the set of priors as $\Pi = \left\{1 - \frac{u}{2}, 1, 1 + \frac{u}{2}\right\}$ such that the amount of manager's ambiguity $\|\Pi\| = u$. For this simulation, I fix the degree of the manager's ambiguity aversion as $\alpha = 8$, and set $\tau = 0.35$ and $k = 0.4$. I change only the amount of uncertainty $\|\Pi\|$ from 0 to 0.8, and numerically determine the optimal leverage decisions of the ambiguity-aversion augmented model with varying $\|\Pi\|$. Figure 1.2 on page 38 presents the numerical results. Point B in the figure represents the optimal leverage decision when the manager faces no uncertainty about the distribution of cash flow. As can be seen, the manager who faces the amount of uncertainty $\|\Pi\| = 0.3$, 30% dispersion from the reference mean ($\theta_2 = 1$), takes substantially less debt (0.61) than the manager who faces no uncertainty (0.83).

## 1.D   Risk-Averse but Ambiguity-Neutral Manager: Numerical Examples

I first characterize a manager's risk aversion as the constant absolute risk aversion function (CARA) such as:

$$u(x) = 1 - e^{-\rho x} \qquad \text{where } \rho > 0\,, \tag{1.32}$$

where the coefficient of absolute risk aversion is $-u''(x)/u'(x) = \rho$. Besides, I set the exogenous parameters as $\tau = 0.35$ and $k = 0.4$. The risk-averse but ambiguity-neutral manager's optimal amount of debt is computed using the first-order condition (1.17), and the leverage ratio is computed using (1.28). The three lines (2), (3), and (4) 1.3 represent

the choice of optimal leverages of the risk-averse but ambiguity-neutral manager when her relative risk coefficient increases from 0.5 to 2.

I turn to compute the optimal leverage chosen by the ambiguity-averse but risk-neutral manager. The three lines (5), (6), and (7) in Figure 1.3 represent the results. I assume the manager displays the constant absolute ambiguity aversion as (1.9). The coefficient of ambiguity aversion increases from $\alpha = 3$ to $\alpha = 6$. The computation procedures are the same as the first simulation of Section 1.C.

## 1.E   Measuring Asset Volatility

Since we can rarely observe the asset value, let alone the asset volatility, of a firm, I estimate the asset volatility of a firm implied from the volatility of observed equity return process. Specifically, following Faulkender and Petersen (2006), I estimate the asset volatility by multiplying the equity volatility by the equity-asset-ratio. The precise formula for asset volatility at time $t$ is:

$$\sigma_{V_t} = \sqrt{\left(\frac{E_t}{V_t}\right)^2 \sigma_{E_t}^2 + \left(\frac{D_t}{V_t}\right)^2 \sigma_{D_t}^2 + 2\left(\frac{D_t}{V_t}\right)\left(\frac{E_t}{V_t}\right) \rho\, \sigma_{D_t}\sigma_{E_t}}, \qquad (1.33)$$

where $V$ is the market value of asset, $E$ the market value of equity, $D$ the total debt, $\sigma_D$ the volatility of debt, and $\rho$ the correlation coefficient between debt and equity. Assuming $\sigma_D = 0$, I have:

$$\sigma_{V_t} = \frac{E_t}{V_t}\sigma_{E_t}. \qquad (1.34)$$

Comparing with (1.33), the approximation (1.34) understates the true asset volatility except for the all equity firm.[21] Also, the magnitude of error increases with the debt-to-asset ratio. However, the difference-in-difference analysis also relies on the change of control variables. Hence, this type of measurement error will be mitigated.

The GARCH model is used to estimate the time-varying equity volatility. That is, tomorrow's volatility is predicted to be a weighted average of the (i) long run average variance, (ii) today's volatility (forecasted from yesterday), and (iii) the news effect (today's squared

---

[21] Alternatively, the Merton formula provides the approximation of volatility of asset such as:

$$\sigma_{V_t} = \frac{\partial V/V}{\partial E/E}\sigma_{E_t} = \frac{1}{\Delta(\sigma_{V_t})}\frac{E}{V}\sigma_{E_t},$$

where $\Delta(V_t)$ is the option's delta. Faulkender and Petersen (2006) report that this alternative does not affect the results when running capital structure panel regression (Faulkender and Petersen, 2006, p. 60).

equity return).

$$r_{it} = \epsilon_{it}\sqrt{\sigma_{E,it}} \qquad \epsilon_{it} \sim N(0,1), \tag{1.35}$$

$$\sigma_{E,it} = \underbrace{c_i}_{(i)} + b_i \underbrace{\sigma_{E,it-1}}_{(ii)} + a_i \underbrace{r_{it-1}^2}_{(iii)}, \tag{1.36}$$

so that a firm $i$'s conditional equity volatility at time $t$ given information $t-1$,

$$r_{it}|\mathcal{F}_{t-1} \sim N(0, \sigma_{E,it}). \tag{1.37}$$

The GARCH-based volatility estimate is a more accurate time-varying risk measure compared to the historical standard deviation approach. According to (1.35), the GARCH model emphasizes a surprise close to the event for study, therefore it incorporates the realized, as well as the forward-looking risk of an asset when the 1982-VRA news arrives.
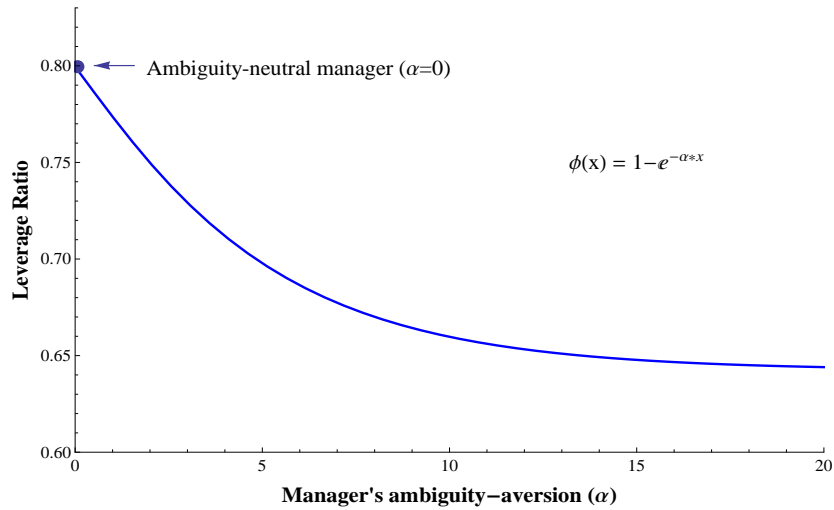
Point A denotes the leverage choice taken by a ambiguity-neutral manager ($\alpha = 0$). The manager is assumed to display the constant absoluteref ambiguity aversion characterized in (1.9). The set of priors and associated beliefs are set as $\Pi = \{0.7, 1, 1.3\}$ and $\{\mu_i\}_{i=1}^3 = 1/3$. The common and known standard deviation is set as $\sigma = 0.3$, the marginal tax rate $\tau = 0.35$, and the proportional deadweight cost $k = 0.4$. The leverage ratio is defined as $B^*/(E(B^*) + B^*)$ where $E(B^*)$ is the market value of equity at $B^*$ computed using (1.3). The computation details to generate this figure are in Appendix 1.C.

**Figure 1.1 Optimal leverage ratio when ambiguity aversion parameter $\alpha$ increases**
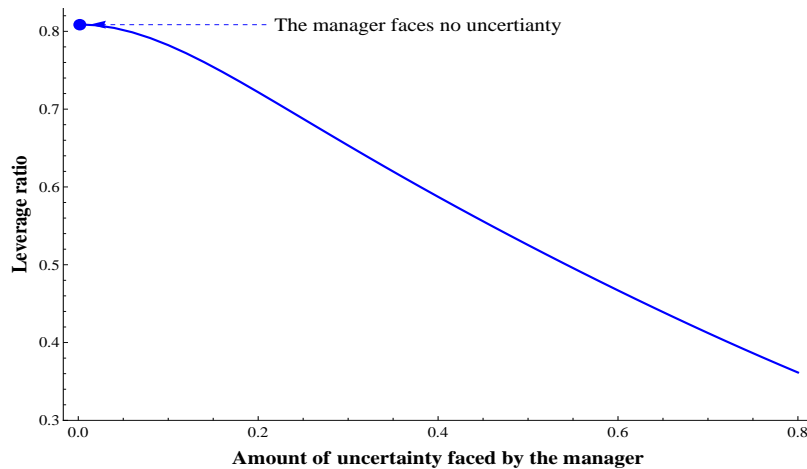


This figure illustrates the optimal leverage ratios when changing the amount of uncertainty perceived by the manager. As the amount of uncertainty increases, the optimal leverage decreases. To generate this figure, the manager's ambiguity aversion parameter is fixed as $\alpha = 8$. The other parameters are set as $\tau = 0.35$, $k = 0.4$, and $\sigma = 0.3$.

**Figure 1.2 Optimal leverage ratio when the amount of uncertainty increases.**
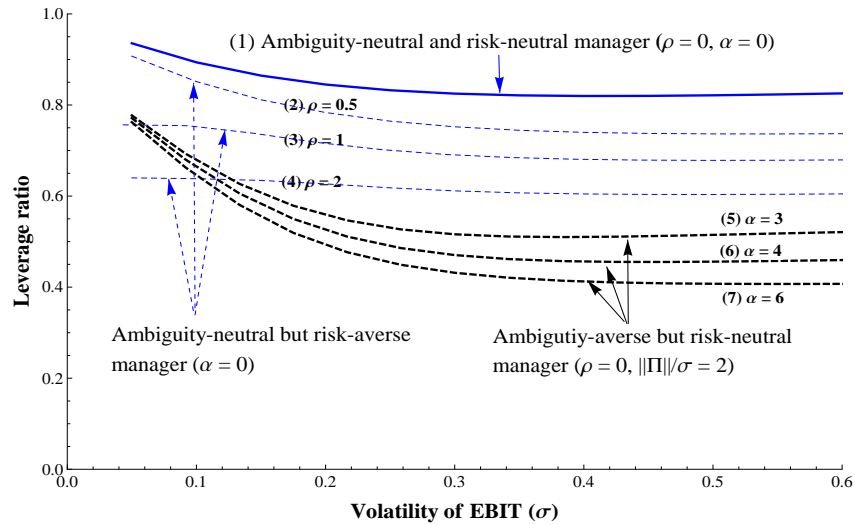
This figure presents the comparative statics of optimal leverage with respect to the risk of a firm's cash flow. The top solid line (1) represents the case where a manager is risk neutral and ambiguity neutral. The three dotted lines (2) - (4) represent the optimal leverage decisions when the manager is risk-averse but ambiguity-neutral when the degree of manager's risk aversion increases from $\rho = 0.5$ to $\rho = 2$. The black dashed lines (5), (6), and (7) denote the optimal leverages when the manager is ambiguity-averse but risk-neutral ($[\alpha = 3$ for line (4), $\alpha = 4$ for line (5), and $\alpha = 6$ for line (6)]). Note that the amount of uncertainty faced by the managers is set as amount of uncertaity$/\sigma = 2$, the marginal corporate tax rate and proportional bankruptcy cost as $\tau = 0.35$ and $k = 0.4$ for this figure.

**Figure 1.3 Comparison of the optimal leverage between the ambiguity-neutral but risk-averse manager and the ambiguity-averse but risk-neutral manager**

**Table 1.1 Variable Definitions**

| Variables | Definition |
|---|---|
| Book Equity | Stockholders Equity + Deferred Taxes and Investment Tax Credit - Preferred Stock (PSTKQ) if (PSTKQ) missing then Preferred Stock Redemption Value (PSTKRQ) |
| Total Debt | Long term debt + Short term debt |
| Market Equity | Common Shares Outstanding * Closing Price |
| Market Leverage | Total Debt / Market Value of Asset |
| Book Leverage | Total Debt / Total Asset |
| Market-to-Book | (Market Equity + Total Debt + Preferred stock value - Deferred taxes and investment tax credits) / Total Asset |
| Firm Size | log(Sales [millions of dollars (henceforth, MM$)]) |
| Tangibility | Plant Property and Equipment / Total Asset |
| Profitability | Earnings Before Interest, Taxes, Depreciation and Amortization / Total Asset |
| Marginal Tax Rate (MTR) | Simulated Marginal Tax Rate available in Graham (1996) |
| Price-cost margin | (Sales - Cost of Goods Sold - Selling Expense)/ Total Asset |

**Table 1.2 Comparison of the Firm-specific Control Variables between the Treatment (U.S. Steel Industry) and Control Groups Before and After Propensity Score Matching**

This table presents the results of two sample tests to compare the means of the treatment and matched control groups before and after the propensity score matching described in Section 1.3.2.3. Prior to matching, the means of each pretreatment characteristic variable of the treatment and matched control group are statistically significantly different, based on the sufficiently large $t$-statistics. However, after matching, they are statistically indistinguishable based on the sufficiently small $t$-statistics. The standard bias of each variable measures a standardized differences in mean: the difference in means of each control variable of the treated and control group divided by the standard deviation of the treatment group, $(\overline{X}_T - \overline{X}_C)/\overline{\sigma}_T$. The bottom table presents the result of Wilcoxon rank-sum test of the null hypothesis whether the distributions of the estimated propensity scores of the treatment and control group are same after matching.

| | Pre-Matching | | | | | Post-Matching | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | | Std. Bias | Two sample $t$-test | | Mean | | Std. Bias | Two sample $t$-test | |
| Variables | Treated | Control | $(\overline{X}_T - \overline{X}_C)/\overline{\sigma}_T$ | $t$-stats | $p$-val | Treated | Control | $(\overline{X}_T - \overline{X}_C)/\overline{\sigma}_T$ | $t$-stats | $p$-val |
| Market Leverage | 0.44 | 0.29 | 0.62 | 3.86 | 0.00 | 0.44 | 0.43 | 0.02 | 0.11 | 0.92 |
| Book Leverage | 0.26 | 0.25 | 0.06 | 0.34 | 0.73 | 0.26 | 0.27 | -0.05 | -0.24 | 0.81 |
| Cash/Asset | 0.08 | 0.12 | -0.32 | -1.68 | 0.09 | 0.08 | 0.08 | -0.03 | -0.15 | 0.89 |
| Tangibility | 0.45 | 0.36 | 0.50 | 2.74 | 0.01 | 0.45 | 0.37 | 0.46 | 2.00 | 0.05 |
| Firm Size | 5.35 | 4.26 | 0.55 | 2.99 | 0.00 | 5.35 | 4.91 | 0.22 | 0.99 | 0.33 |
| Mkt-to-Book | 0.62 | 1.33 | -0.63 | -2.81 | 0.01 | 0.62 | 0.62 | 0.01 | 0.11 | 0.91 |
| Profitability | 0.04 | -0.02 | 0.07 | 0.32 | 0.75 | 0.04 | 0.08 | -0.05 | -1.62 | 0.11 |
| vol(Asset) | 0.23 | 0.36 | -0.78 | -3.91 | 0.00 | 0.23 | 0.24 | -0.04 | -0.20 | 0.84 |
| Annual Stock Ret. | -0.12 | 0.23 | -0.67 | -3.66 | 0.00 | -0.12 | -0.12 | 0.00 | 0.01 | 0.99 |
| Marginal Tax rate | 0.40 | 0.41 | -0.10 | -0.13 | 0.90 | 0.40 | 0.42 | -0.20 | -1.04 | 0.30 |

Two-sample Wilcoxon rank-sum (Mann-Whitney) test
$H_0$: Distributions of estimated propensity score of the treatment and the matched control group are same
$p$-val $= 0.719$
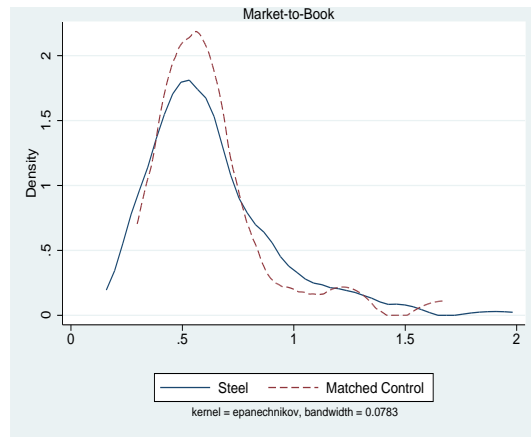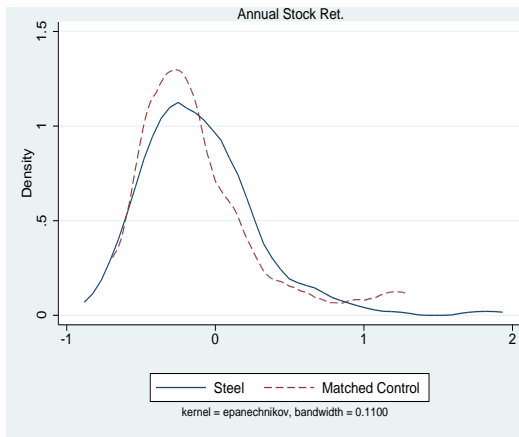
**(a)** Market-to-Book Ratio  **(b)** Annual Stock Return  **(c)** Market Leverage
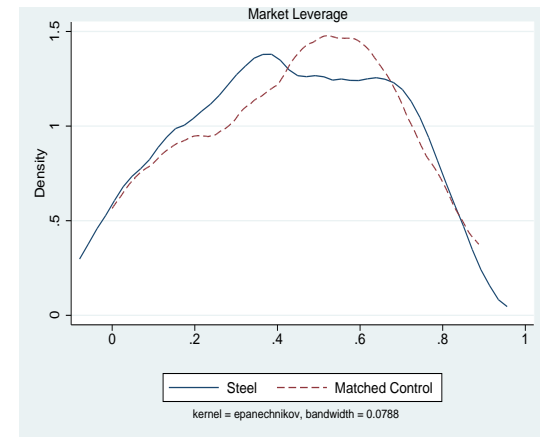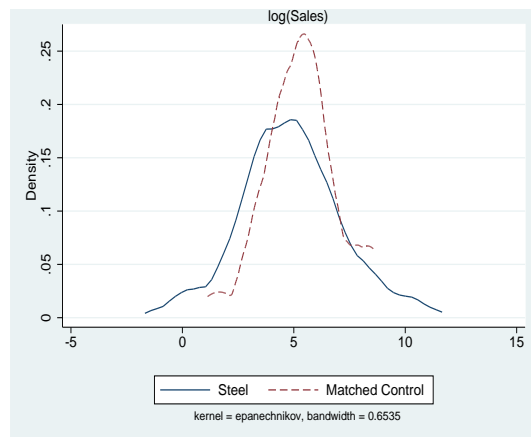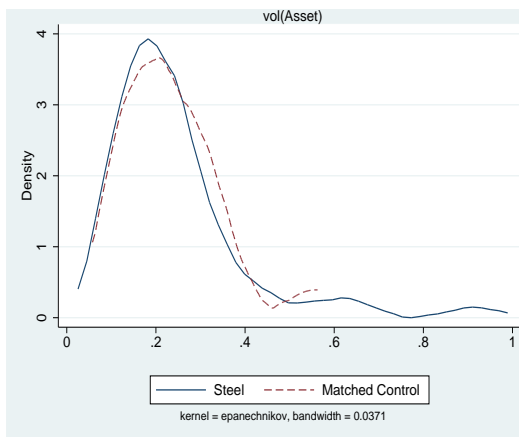
**(d)** Firm Size  **(e)** Asset Volatility  **(f)** Profitability
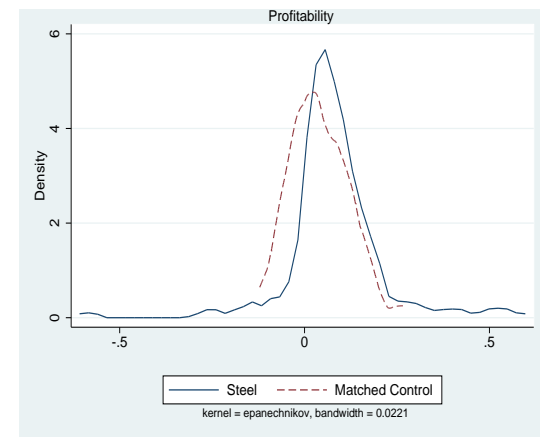
The solid lines represent the estimated Epanechnikov kernel densities of each variable of the treatment group (U.S. steel manufacturers), the dashed lines denote those of the matched control group.

**Figure 1.4 Post-matching Balance of Firm Characteristics Variables**

**Table 1.3 Summary Statistics of the Firm-specific Control Variables of Treatment and Matched Control Groups in the Pre- and Post-VRA Period**

Panel A represents the summary statistics of the firm-specific control variables of the treatment and matched control group in the pre-VRA period (1978-1982). To construct the matched control group, I use the standard propensity score matching described in Section 1.3.2.3. Panel B represents the summary statistics of the firm-specific control variables of the treatment and matched control groups in the post-VRA period (1983-1987). The definitions of variables are in Table 1.1.

Panel A: Pre-VRA Period (1978-1982)

| Firm Characteristics | Treatment | | | Matched Control | | |
|---|---|---|---|---|---|---|
| | Mean | Std | Median | Mean | Std | Median |
| Profitability | 0.143 | 0.095 | 0.137 | 0.138 | 0.086 | 0.140 |
| Mkt-to-Book | 0.685 | 0.348 | 0.623 | 0.740 | 0.522 | 0.628 |
| Tangibility | 0.421 | 0.136 | 0.430 | 0.351 | 0.200 | 0.336 |
| Firm Size | 5.579 | 1.523 | 5.510 | 5.029 | 2.114 | 4.869 |
| vol(Asset) | 0.247 | 0.113 | 0.244 | 0.249 | 0.131 | 0.225 |
| Ann. Stock Ret | 0.128 | 0.416 | 0.067 | 0.147 | 0.458 | 0.078 |
| MTR (before int.) | 0.429 | 0.056 | 0.460 | 0.436 | 0.064 | 0.460 |

Panel B: Post-VRA Period (1983-1987)

| Firm Characteristics | Treatment | | | Matched Control | | |
|---|---|---|---|---|---|---|
| | Mean | Std | Median | Mean | Std | Median |
| Profitability | 0.081 | 0.090 | 0.094 | 0.080 | 0.180 | 0.102 |
| Mk-to-Book | 0.732 | 0.299 | 0.663 | 0.782 | 0.320 | 0.714 |
| Tangibility | 0.415 | 0.156 | 0.417 | 0.362 | 0.212 | 0.346 |
| Firm Size | 5.506 | 1.791 | 5.463 | 4.887 | 2.362 | 4.636 |
| vol(Asset) | 0.255 | 0.116 | 0.247 | 0.264 | 0.145 | 0.232 |
| Ann. Stock Ret | 0.174 | 0.483 | 0.128 | 0.171 | 0.465 | 0.105 |
| MTR (before int.) | 0.362 | 0.112 | 0.390 | 0.392 | 0.112 | 0.453 |

**Table 1.4 Necessary Condition for Parallel-Trends Assumption in the Pre-VRA Period**

This table presents the results of checking the necessary condition of the parallel-trends assumption. Columns (1) and (2) represent the means and standard deviations of leverage growth rate of the treatment and matched control group in the pre-VRA period (1978-1982). Column (3) represents the results of two sample t-test to explore whether the means of leverage growth rate of the treatment and matched control group are statistically different. Column (4) reports the results of Wilcoxon rank-sum test of the null hypothesis if two samples are drawn from the same distributions.

| | (1) Treatment | | (2) Matched Control | | (3) Two sample test | | | (4) Ranksum |
|---|---|---|---|---|---|---|---|---|
| Growth rate | Mean | Std | Mean | Std | Diff | t-stats | p-val | p-val |
| ML (1978-1979) | 0.198 | 0.686 | 0.037 | 0.534 | -0.161 | -1.453 | 0.148 | 0.214 |
| ML (1979-1980) | 0.206 | 1.501 | 0.249 | 2.088 | 0.044 | 0.115 | 0.908 | 0.745 |
| ML (1980-1981) | 0.040 | 0.391 | 0.324 | 1.569 | 0.284 | 1.074 | 0.284 | 0.189 |
| ML (1981-1982) | 0.605 | 2.141 | 0.388 | 1.179 | -0.217 | -0.830 | 0.408 | 0.654 |
| | | | | | | | | |
| BL (1978-1979) | 0.153 | 0.639 | 0.066 | 0.587 | -0.087 | -0.754 | 0.452 | 0.822 |
| BL (1979-1980) | 0.291 | 1.593 | 0.265 | 2.016 | -0.026 | -0.071 | 0.944 | 0.552 |
| BL (1980-1981) | -0.036 | 0.247 | 0.277 | 2.344 | 0.313 | 0.798 | 0.426 | 0.246 |
| BL (1981-1982) | 0.358 | 1.303 | 0.178 | 0.812 | -0.180 | -1.055 | 0.293 | 0.118 |

Checking Parallel-Trends (Market Leverage)
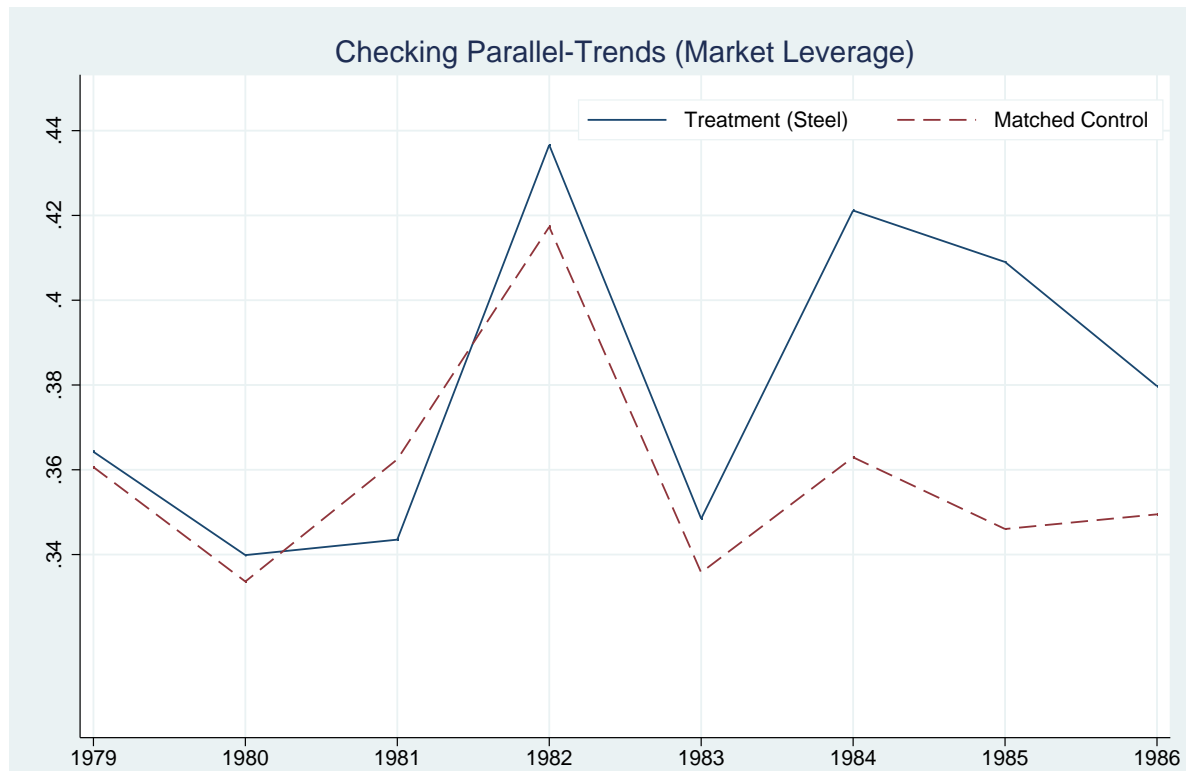
This figure provides the graphical examination of parallel-trends assumption for market leverage. If the parallel-trends holds, the pre-treatment trends across the treatment and control groups are parallel. The figure illustrates the market-leverage trend of the treatment group (steel industry) and that of the matched control group using the standard four nearest propensity score matching. The pre-VRA period is the period between 1978 and 1982 and the post-VRA period is the period between 1983 and 1987. The pre-VRA trends across the two groups seem parallel, whereas the market-leverage trend of the treatment seems to diverge from that of the matched control group in the post-VRA era. It implies that in the pre-VRA period, the parallel-trends assumption is met.

**Figure 1.5 The Pre- and Post-trends of the Treatment and Matched Control Groups**

**Table 1.5 Difference-in-Differences with the Matched Sample**

This table presents the main results of difference-in-difference (DID) analysis with the matched sample specified in (1.14). The outcome variable is market and book leverage. The treatment group is the U.S. steel industry. The matched control firms are constructed using the standard propensity score matching. Each firm in the steel industry is matched up to nearest four non-steel firms in terms of its estimated propensity score. The post-treatment variable $D(i,t) = 1$ if a firm belongs to the U.S. steel industry (the treatment group) and is in the post-VRA period (1983-1987). The estimated coefficient of $D(i,t)$ (i.e., the DID estimate) represents the estimated causal effect of resolution of uncertainty on steel firms' leverage relative to matched control firms. The definitions of other variables are provided in Table 1.1. All regressions include firm and year fixed effects. Standard errors are computed with clustering at the firm-level, thus robust to heteroscedasticity and within-firm serial correlation. The $t$-statistics of the coefficient estimates are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by "*", "**", and "***", respectively.

| | Model 1 Market Lev. | Model 2 Book Lev. | Model 3 Market Lev. | Model 4 Book Lev. |
|---|---|---|---|---|
| D(i,t) | 0.055*** | 0.032** | 0.046** | 0.028* |
| | (3.034) | (2.158) | (2.510) | (1.921) |
| Profitability | -0.116 | -0.045 | -0.107 | -0.042 |
| | (-1.399) | (-0.651) | (-1.194) | (-0.634) |
| Tangibility | 0.299*** | 0.222*** | 0.252*** | 0.204*** |
| | (3.979) | (3.471) | (3.635) | (3.397) |
| Firm Size | 0.033 | 0.022 | 0.048** | 0.029* |
| | (1.585) | (1.363) | (2.581) | (1.741) |
| Market-to-Book | -0.010 | 0.094** | -0.003 | 0.096*** |
| | (-0.384) | (2.614) | (-0.160) | (2.831) |
| Med. Analyst For. | -0.021*** | -0.006** | -0.020*** | -0.005 |
| | (-4.208) | (-2.008) | (-4.000) | (-1.539) |
| vol(Asset) | -1.281*** | -0.817*** | -1.374*** | -0.905*** |
| | (-8.099) | (-7.864) | (-9.599) | (-8.429) |
| Annual Stock Ret. | -0.014 | 0.005 | -0.010 | 0.011 |
| | (-1.338) | (0.568) | (-0.874) | (1.216) |
| Excess Price-cost Mar. | | | 0.102 | 0.100 |
| | | | (1.570) | (1.550) |
| Constant | 0.413*** | 0.165 | 0.351*** | 0.147 |
| | (2.996) | (1.654) | (2.688) | (1.392) |
| Firm fixed effects | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Within $R^2$ | 0.671 | 0.442 | 0.691 | 0.475 |
| Num. of Obs. | 735 | 735 | 693 | 693 |

**Table 1.6 Additional Test (Cash/Asset): Necessary Condition for Parallel-Trends Assumption in the Pre-VRA Period**

This table presents the results of checking the necessary condition of the parallel-trends assumption. Columns (1) and (2) represent the means and standard deviations of cash holding ratio (cash/asset) growth rate of the treatment and matched control group in the pre-VRA period (1978-1982). Column (3) presents the results of two sample t-test to explore whether the means of leverage growth rate of the treatment and matched control group are statistically different. Column (4) reports the results of Wilcoxon rank-sum test of the null hypothesis if two samples are drawn from the same distributions.

| | (1) Treatment | | (2) Matched Control | | (3) Two sample test | | | (4) Ranksum |
|---|---|---|---|---|---|---|---|---|
| Growth rate | Mean | Std | Mean | Std | Diff | t-stats | p-val | p-val |
| Cash (1978-1979) | 0.175 | 0.984 | 0.466 | 2.055 | 0.291 | 0.821 | 0.413 | 0.773 |
| Cash (1979-1980) | 1.435 | 5.453 | 0.367 | 2.302 | -1.068 | -1.741 | 0.084 | 0.691 |
| Cash (1980-1981) | 0.392 | 1.429 | 0.170 | 1.262 | -0.222 | -0.917 | 0.361 | 0.322 |
| Cash (1981-1982) | 0.799 | 2.069 | 0.279 | 1.474 | -0.520 | -1.782 | 0.076 | 0.235 |

**Table 1.7 Additional Test (Cash/Asset): Difference-in-Differences with the Matched Sample**

This table presents the additional results of difference-in-difference (DID) analysis with the matched sample specified in (1.14). The outcome variable is cash/asset ratio. The treatment group is the U.S. steel industry. The matched control firms are constructed using the standard propensity score matching. Each firm in the steel industry is matched up to nearest four non-steel firms in terms of its estimated propensity score. The post-treatment variable $D(i,t) = 1$ if a firm belongs to the steel producing industry (the treatment group) and is in the post-VRA period (1983-1987). The definitions of other variables are provided in Table 1.1. The estimated coefficient of $D(i,t)$ (i.e., the DID estimate) represents the causal effect of resolution of uncertainty on cash holding ratio. All regressions include firm and year fixed effects. Standard errors are computed with clustering at the firm-level, thus robust to heteroscedasticity and within-firm serial correlation. The $t$-statistics of the coefficient estimates are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by "*", "**", and "***", respectively.

|  | Model 1 Cash/Asset | Model 2 Cash/Asset |
|---|---|---|
| D(i,t) | -0.022* | -0.022 |
|  | (-1.707) | (-1.635) |
| Profitability | 0.058 | 0.113 |
|  | (0.772) | (1.181) |
| Tangibility | -0.334*** | -0.340*** |
|  | (-5.497) | (-5.150) |
| Firm Size | -0.003 | -0.004 |
|  | (-0.232) | (-0.239) |
| Market-to-Book | -0.036** | -0.036* |
|  | (-2.072) | (-1.918) |
| Med. Analyst For. | -0.006* | -0.006** |
|  | (-1.972) | (-2.001) |
| vol(Asset) | 0.085* | 0.097* |
|  | (1.730) | (1.707) |
| Annual Stock Ret. | 0.008 | 0.009 |
|  | (1.205) | (1.278) |
| Excess Price-cost Mar. |  | 0.039 |
|  |  | (0.881) |
| Constant | 0.238*** | 0.232** |
|  | (2.620) | (2.376) |
| Firm fixed effects | Yes | Yes |
| Year fixed effects | Yes | Yes |
| Within $R^2$ | 0.203 | 0.207 |
| Num. of Obs. | 734 | 692 |

**Table 1.8 Collapsing the Data into Two Periods (à la Bertrand et al., 2004)**

This table presents the results of difference-in-differences (DID) estimator ignoring time series information proposed by Bertrand et al. (2004). Following Bertrand et al., I time-average the data before and after the 1982 VRA and run analysis (1.14) on this averaged outcome variable in a panel of length 2. That is, the DID model (1.14) becomes $\Delta Y_i = \beta_0 + \beta D_i + \gamma' \Delta X_i + \Delta u_i$. $\Delta$ represents the change between pre- and post-variables. The treatment $D(i) = 1$ if a firm belongs to the U.S. steel industry. $\beta$ represent the DID estimator: the causal effect of resolution of uncertainty on the change of leverage. As Bertrand et al. (2004) suggest, this is most conservative way robust to the within-firm serial correlation of errors. The $t$-statistics are reported in parentheses. All standard errors are robust to heteroscedasticity. Statistical significance at the 10%, 5%, and 1% levels are denoted by "*", "**", and "***", respectively.

|  | Model 1<br>$\Delta$ Market Lev. | Model 2<br>$\Delta$ Book Lev. | Model 3<br>$\Delta$ Market Lev. | Model 4<br>$\Delta$ Book Lev. |
|---|---|---|---|---|
| D(i) | 0.057*** | 0.029 | 0.061*** | 0.036* |
|  | (2.642) | (1.562) | (2.721) | (1.889) |
| $\Delta$ Profitability | -0.018 | 0.126 | 0.377 | 0.420 |
|  | (-0.070) | (0.566) | (0.959) | (1.137) |
| $\Delta$ Market-to-Book | 0.029 | 0.175*** | -0.004 | 0.164*** |
|  | (0.544) | (3.517) | (-0.073) | (3.137) |
| $\Delta$ Tangibility | 0.523*** | 0.374*** | 0.480*** | 0.381*** |
|  | (3.584) | (3.014) | (3.455) | (3.256) |
| $\Delta$ Firm Size | -0.011 | -0.021 | 0.008 | -0.023 |
|  | (-0.301) | (-0.729) | (0.236) | (-0.691) |
| $\Delta$ Med. Anal. Forecast | -0.032*** | -0.016** | -0.034*** | -0.016** |
|  | (-3.497) | (-2.511) | (-3.676) | (-2.335) |
| $\Delta$ vol(Asset) | -1.478*** | -1.113*** | -1.464*** | -1.151*** |
|  | (-6.846) | (-8.956) | (-5.926) | (-7.666) |
| $\Delta$ Annual Stock Ret. | 0.048 | 0.083*** | 0.045 | 0.082*** |
|  | (1.177) | (2.798) | (1.132) | (2.819) |
| $\Delta$ Excess Price-cost Mar. |  |  | 0.418 | 0.429 |
|  |  |  | (1.213) | (1.244) |
| Constant | -0.051*** | -0.018 | -0.039** | -0.011 |
|  | (-3.129) | (-1.465) | (-2.175) | (-0.723) |
| Adj. $R^2$ | 0.690 | 0.595 | 0.679 | 0.604 |
| Num. of Obs. | 93 | 93 | 87 | 87 |

**Table 1.9 Difference-in-Difference with the Placebo Shocks at 1979, 1993, and 2001**

This table presents the results of tests with placebo-VRAs had they taken place in 1979, 1993 or 2001 instead of 1982. In each test, I use the difference-in-difference methodology with the propensity-score based matched sample the same as Section 1.3.2.3. In all tests, none of the DID estimates (i.e., the estimated coefficients of $D(i,t)$) is statistically significantly positive, observing the sufficiently small $t$-statistics. Standard errors are computed with clustering at the firm-level, thus robust to heteroscedasticity and within-firm serial correlation. The $t$-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by "*", "**", and "***", respectively.

| | 1979 Placebo Shock | | 1993 Placebo Shock | | 2001 Placebo Shock | |
|---|---|---|---|---|---|---|
| | Market Lev. | Book Lev. | Market Lev. | Book Lev. | Market Lev. | Book Lev. |
| D(i,t) | 0.005 | 0.003 | 0.018 | 0.003 | -0.023 | -0.033 |
| | (0.262) | (0.247) | (0.893) | (0.182) | (-0.713) | (-1.497) |
| Profitability | -0.819*** | -0.357*** | 0.020 | 0.035 | -0.053 | -0.144 |
| | (-7.191) | (-3.045) | (0.190) | (0.474) | (-0.307) | (-1.198) |
| Tangibility | 0.080 | 0.163** | 0.136* | 0.098 | 0.024 | 0.166* |
| | (0.999) | (2.032) | (1.750) | (1.599) | (0.240) | (1.887) |
| Firm Size | 0.105*** | 0.079*** | 0.032 | 0.050** | 0.043 | 0.052* |
| | (4.120) | (3.084) | (1.461) | (2.506) | (1.205) | (1.806) |
| Market-to-Book | -0.050** | 0.031 | -0.070*** | 0.010 | -0.062* | 0.062** |
| | (-2.200) | (1.454) | (-3.738) | (0.684) | (-1.925) | (2.417) |
| Med. Analyst For. | 0.000 | -0.000** | -0.020*** | -0.012** | -0.009** | -0.003 |
| | (0.055) | (-2.356) | (-3.610) | (-2.085) | (-2.178) | (-1.101) |
| vol(Asset) | -0.879*** | -0.499*** | -0.838*** | -0.562*** | -0.833*** | -0.478*** |
| | (-6.056) | (-5.746) | (-6.855) | (-5.708) | (-8.808) | (-6.508) |
| Annual Stock Ret. | -0.009 | 0.016** | -0.001 | 0.017* | -0.021* | 0.008 |
| | (-0.843) | (2.400) | (-0.096) | (1.930) | (-1.719) | (0.760) |
| Constant | 0.067 | -0.150 | 0.472*** | 0.127 | 0.385 | 0.004 |
| | (0.403) | (-1.084) | (3.379) | (1.083) | (1.661) | (0.019) |
| Firm & Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Within $R^2$ | 0.641 | 0.399 | 0.585 | 0.380 | 0.690 | 0.447 |
| Num. of Obs. | 691 | 691 | 950 | 950 | 436 | 436 |

**Table 1.10 Difference-in-Difference with Alternative Matching Methods**

This table presents the results of difference-in-difference with the matched sample using alternative matching methods: (I) Mahalanobis-metric matching with a caliper 0.05, and (II) propensity score matching with the different set of matching variables. As before, the treatment group is the U.S. steel industry. The outcome variables are market and book leverage ratio. Case (I) Mahalanobis-metric matching with a caliper 0.05: I take each firm in the steel industry and match it up to nearest four non-steel firms in terms of Mahalanobis-metric instead of the propensity score distance. The firm characteristics that I match on are the same as Section 1.3.2.3. Case (II) the propensity score matching with the different set of matching variables: I match on the firm size, tangibility, profitability, and market-to-book ratio. The unreported results suggest that both alternative matching methods are successful in balancing the pretreatment firm characteristics. All standard errors are computed with clustering at the firm-level. The $t$-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by "*", "**", and "***", respectively.

| | Alternative Matching (I) | | | Alternative Matching (II) | | |
|---|---|---|---|---|---|---|
| | Market Leverage | Book Leverage | Cash/Asset | Market Leverage | Book Leverage | Cash/Asset |
| D(i,t) | 0.050*** | 0.027** | -0.021* | 0.058*** | 0.029* | -0.021* |
| | (2.609) | (1.986) | (-1.707) | (3.086) | (1.909) | (-1.717) |
| Profitability | -0.178*** | -0.075 | 0.039 | -0.181** | -0.170*** | 0.145*** |
| | (-2.991) | (-1.146) | (0.819) | (-2.014) | (-2.935) | (2.764) |
| Tangibility | 0.114 | 0.029 | -0.402*** | 0.167** | 0.091 | -0.373*** |
| | (1.472) | (0.347) | (-5.694) | (2.165) | (1.375) | (-4.256) |
| Firm Size | 0.040 | 0.039* | -0.029 | 0.021 | 0.031* | 0.015 |
| | (1.239) | (1.770) | (-1.624) | (1.029) | (1.793) | (0.899) |
| Market-to-Book | -0.059* | 0.096** | 0.011 | -0.065*** | 0.086*** | -0.026 |
| | (-1.955) | (2.543) | (0.538) | (-2.834) | (3.954) | (-1.155) |
| vol(Asset) | -0.768*** | -0.470*** | -0.109 | -1.229*** | -0.634*** | 0.045 |
| | (-3.520) | (-3.162) | (-1.567) | (-12.112) | (-7.992) | (0.715) |
| Annual Stock Ret. | -0.026*** | -0.001 | 0.011** | -0.002 | 0.001 | 0.012* |
| | (-2.814) | (-0.115) | (2.260) | (-0.251) | (0.192) | (1.669) |
| Med. Analyst For. | -0.032*** | -0.017*** | 0.000 | -0.002 | -0.001 | -0.001 |
| | (-5.142) | (-3.874) | (0.080) | (-0.333) | (-0.208) | (-0.460) |
| Constant | 0.384* | 0.074 | 0.440*** | 0.555*** | 0.153 | 0.104 |
| | (1.652) | (0.474) | (4.005) | (4.615) | (1.415) | (1.146) |
| Firm & Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Within-in $R^2$ | 0.614 | 0.285 | 0.223 | 0.693 | 0.421 | 0.237 |
| Number of Obs. | 899 | 899 | 898 | 682 | 682 | 681 |

# Chapter 2

# A New Measure of Knightian Uncertainty

## 2.1 Introduction

Theoretical work on decision making in the presence of Knightian uncertainty is abundant (for example, see Gilboa and Schmeidler (1989) and the subsequent literature). However, to the best of my knowledge, efforts to estimate firm-level uncertainty directly from the data seem limited. As proxies for uncertainty, researchers often use measures such as the analyst forecast dispersion and the earnings response coefficient. One dissatisfaction with these measures is that the data required to compute them are not always reliably available.

As a new measure of the amount of uncertainty faced by a firm, I use the statistical distance between a mixture of normal distributions (i.e., a weighted sum of normal component distributions) and a single normal distribution. As in Chapter 1, in my model, the Knightian manager is unsure about which one of the three normal distributions is the true distribution of the firm's profitability in the next period. Instead, she attaches her own subjective belief (weight) to each distribution being the true distribution.

I consider the realized profitability of a firm over time. First, I fit a single normal distribution to the sample of observed profitability. This is the benchmark case where the manager of the firm is absolutely certain about the distribution of its future profitability. Next, I fit a mixture of three normal distributions to the same data. Finally, I compute the Jensen-Shannon distance (henceforth, JS distance) between the mixture of the three distributions and the single normal distribution. I measure the amount of uncertainty perceived by the firm by this statistical distance. The JS distance is conceptually similar to the

Kullback-Leibler divergence, with the notable difference that it is a finite metric (Pollard, 2002).

I then present empirical evidence that suggests the amount of uncertainty perceived by a firm (i.e., the JS distance) is an important determinant of leverage. I run between-group regressions of leverage on the JS distance (i.e., the firm-specific amount of uncertainty), along with other firm-specific controls and industry fixed effects. I find a statistically significant and economically sizable negative effect of the firm's uncertainty on leverage. Consider a median firm in the sample whose JS distance is 0.253. If this firm's JS distance increases by one standard deviation (0.14), its market leverage decreases by approximately $0.14 \times 0.051$ or 0.71%. Scaling by the mean market leverage in the sample (0.2), a one standard deviation increase in the firm's JS distance decreases its market leverage by approximately 3.6%. The same procedure leads to about a 3% fall in book leverage. The results are economically sizable and statistically significant. The paper therefore adds to what we know of the empirical determinants of a firm's leverage (e.g., Titman and Wessels, 1988; Rajan and Zingales, 1995; Lemmon et al., 2008).

Strebulaev and Yang (2012) document a zero-leverage puzzle: between 1962 and 2006, about 10% of public non-financial firms take on almost zero debt, and 32% have zero or negative net debt. They also point out that the zero-leverage puzzle is not a short-term fad. Can uncertainty provide a reasonable explanation to a firm's extremely low leverage taking tendency? Following Strebulaev and Yang (2012), I classify a firm as an almost zero-leverage firm if its time-averaged leverage is less than 5%. I run a logit regression of the propensity of a firm's taking almost zero leverage on the JS distance. Keeping the other covariates fixed as their mean values, an increase in JS distance from one standard deviation below its mean to one standard deviation above its mean is associated with an increase in the propensity of a firm's taking almost zero leverage by 2.4%; an impact of changing JS distance from the minimum to the maximum is associated with an increase in the propensity of a firm's taking an almost zero leverage by 6.3%.

Section 2.2 provides the crux of the smooth ambiguity aversion model proposed by KMM. The details about the mixture of normal distributions and the firm-level JS distance are also provided. Section 2.3 details the empirical studies. Section 2.5 concludes.

## 2.2 New Firm-specific Measure of Uncertainty

### 2.2.1 New uncertainty measure using a mixture of normal distributions

In Chapter 1, I show that an ambiguity-averse manager translates her physical belief $\mu$ over the set of priors $\Pi$ into an ambiguity-adjusted belief by her degree of ambiguity aversion characterized by the concavity of function $\phi(\cdot)$ (Lemma 1.1 on page 7). As a risk-averse manager faces the same amount of risk as a risk-neutral manager, an ambiguity-averse manager confronts the same amount of uncertainty as an ambiguity-neutral manager (see the details in Lemma 1.1). Therefore, without loss of generality, I assume the manager is ambiguity-neutral, to measure the firm-level amount of uncertainty.

Assume that the set of priors, $\Pi$, is a collection of three normal distributions with a parameter vector $\vartheta_i = (\theta_i, \sigma_i)$ for $i = 1, 2, 3$. Then, $\Pi$ is fully indexed by $\vartheta_i$, and is rewritten as $\Pi = \{N(\theta_i, \sigma_i)\}_{i=1}^3$. Assuming the manager is ambiguity-neutral, (i.e., $\phi(x) = refx$), by Fubini's theorem, her smooth ambiguity-averse preference (1.1) on page 5 becomes:

$$\sum_{i=1}^3 \phi\Big( \int u(x)\, f_N(x; \vartheta_i) dx \Big) \mu(\vartheta_i) = \int u(x)\, \left( \sum_{i=1}^3 \mu_i f_N(x; \vartheta_i) \right) dx\,, \qquad (2.1)$$

where $f_N(x; \vartheta_i)$ is a normal density with parameters $(\theta_i, \sigma_i)$ associated with the normal distribution $N(\theta_i, \sigma_i^2)$ and $\mu_i$ is the manager's associated belief to $N(\theta_i, \sigma_i^2)$ in $\Pi$.

Define the finite mixture density $\Lambda(x)$ associated with the mixture of normal distributions:

$$\Lambda(x) = \sum_{i=1}^3 \mu_i f_N(x; \vartheta_i). \qquad (2.2)$$

If the parameters $\{(\theta_i, \sigma_i), \mu_i\}_{i=1}^3$ of the mixture density $\Lambda(x)$ are estimated from the firm's realized profitability samples, one immediately specifies the manager's set of priors, $\Pi$, and her associated belief $\mu$ over $\Pi$.

Figure 2.1 on the following page presents an example of estimating the parameters of a mixture of three normal distributions. The figure illustrates the histogram of profitability (earnings/asset) of a company in the pharmaceutical industry from 1970 to 2008. Conventionally, one can find the maximum likelihood (ML) estimates of the parameters $\{\mu_i, \theta_i, \sigma_i\}_{i=1}^3$. Dempster et al. (1977) show that under fairly mild regularity conditions, expectation maximization (EM) algorithms more efficiently converge to a local maximum of

the mixture likelihood function (Hamilton, 1994). The details of the EM algorithms used to estimate the parameters of a mixture of normal distributions are in Appendix 2.B on page 65. Using the EM algorithms, I estimate the parameters of the mixture density $\Lambda(x)$ for the pharmaceutical company:

$$\{\theta_i\}_{i=1}^3 = \{-0.25, -0.07, 0.04\}, \{\sigma_i\}_{i=1}^3 = \{0.13, 0.06, 0.02\}, \{\mu\}_{i=1}^3 = \{0.21, 0.55, 0.24\} \tag{2.3}$$
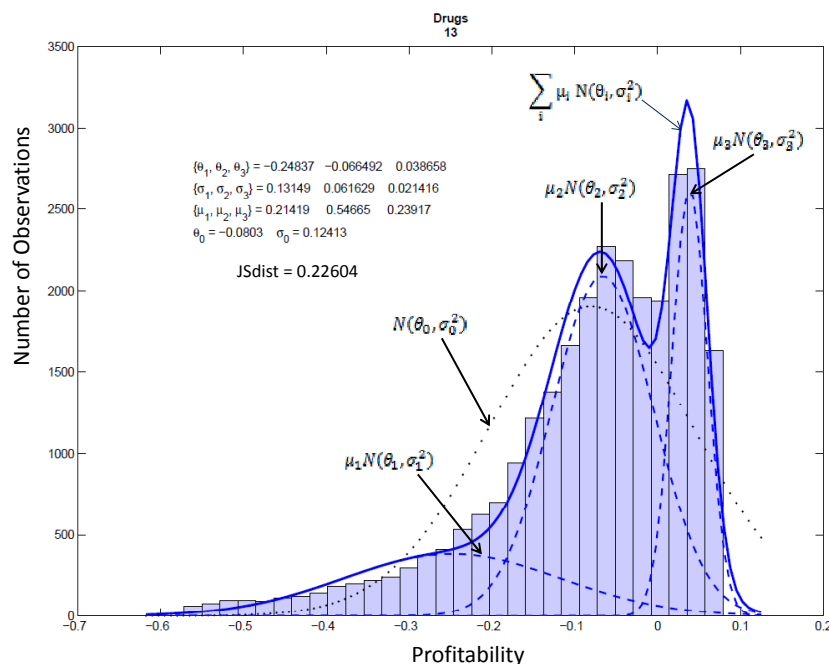


The solid line represents the mixture density $\Lambda(x) = \sum_i^3 \mu_i f_N(x; \theta_i, \sigma_i^2)$ fitted to the sample. The dashed lines represent each of three components weighted by belief $\mu_i \cdot f_N(x; \theta_i, \sigma_i^2)$ for $i = 1, 2, 3$. The dotted line represents the density of the single normal distribution $N(\theta_0, \sigma_0^2)$ fitted to the sample.

**Figure 2.1 Estimation of the mixture of three normal distributions from the realized profitability samples of a representative firm in the pharmaceutical industry between 1970 and 2008**

The manager of the firm faces the empirical distribution as shown in Figure 2.1. The mixture of normal distributions represents the case where the Knightian manager is unsure about which distribution of $\Pi = \{N(\theta_i, \sigma_i^2)\}_{i=1}^3$ is true one for the firm's profitability in the next period; instead, she attaches her own subjective belief $\mu_i$ to each $N(\theta_i, \sigma_i^2)$ being the true distribution. Next, consider the case where the manager is absolutely sure that the firm's next period profitability follows the single normal distribution $N(\theta_0, \sigma_0^2)$. I consider it as the benchmark case where the manager faces no uncertainty.

Next, I measure the firm's amount of uncertainty by the statistical distance between the

single distribution $N(\theta_0, \sigma_0^2)$ and the mixture of three normal distributions $\sum_{i=1}^{3} \mu_i \cdot N(\theta_i, \sigma_i^2)$. The statistical distance between two arbitrary distributions is often measured by the relative entropy (or Kullback-Leibler divergence) from one to another (Pollard 2002; Cover and Thomas 2006). However, the divergence is a quasi-metric that violates the symmetry property. Hence, I use the JS distance to measure the statistical distance between two distributions. The JS distance is the square root of the Jensen-Shannon divergence that is the *symmetrized* Kullback-Leibler divergence. The JS distance is bounded below from 0 and above by 1. The zero JS distance means that two distributions are identical, whereas unity represents complete dissimilarities between two distributions. Figure 2.2 illustrates two examples of JS distance between two distributions. Intuitively, JS distance measures how much the two distributions are similar. Note that large variances of the distributions do not necessarily imply a large JS distance. The details about the formula and properties of the JS distance are in Appendix 2.A on page 65.
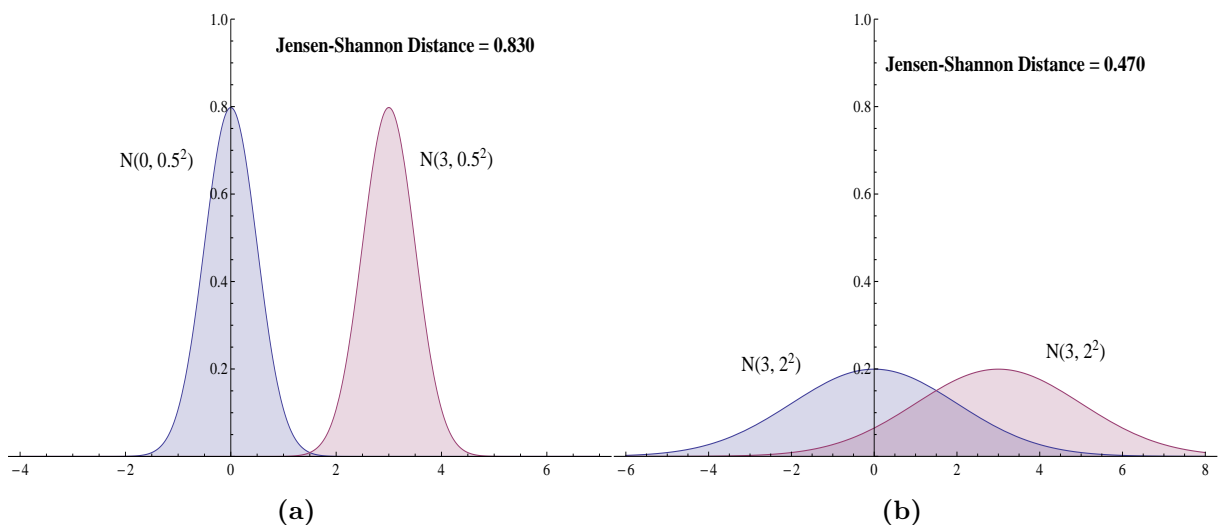


Figure (a) depicts the JS distance (0.8) between two normal distributions $N(0, 0.5^2)$ and $N(3, 0.5^2)$, and Figure (b) shows the JS distance (0.4) between two normal distributions $N(0, 2^2)$ and $N(3, 2^2)$.

**Figure 2.2 Two examples of the JS distance**

Table 2.4 on page 69 presents the JS distances of the 48 Fama-French industries. The "JS dist" column presents the industry-specific amount of uncertainty measured by the JS distance between the mixture of normal distributions and the single normal distribution fitted to the industry-specific profitability samples from 1970 to 2008. The ranking of the industries seems intuitive. The technology industries (e.g., computers, business services, and electrical equipments) face higher uncertainty than the manufacturing industries (e.g., ships and textiles). I argue that the high-tech industries likely employ innovative technologies,

which is riskier as well as more uncertain, than the industries employing mature technology. When a firm employs an innovative technology, Knightian uncertainty is likely due to the lack of historical observations of the profitability of this technology. Accordingly, managers in high-tech industries are presumably unsure of the distribution of the cash flow generated by the innovative technology.

## 2.3   New Uncertainty Measure and Leverage

### 2.3.1   Data

The data set consists of all nonfinancial firm-year observations in the annual Compustat database from 1971 to 2004. I require that all firm-year observations have valid data for book assets. All ratios are winsorized at the upper and lower one percentile to mitigate the effect of outliers and eradicate errors in the data. I obtain analyst median, as well as the highest and lowest forecasts of a firm's next-year earnings per share (EPS) from the IBES database and merge them with the accounting information obtained from the annual Compustat database. The variable definitions are in Table 2.1 on page 67.

### 2.3.2   Pairwise correlations

Table 2.3 on page 68 presents the summary statistics of firm-level JS distance. There are 11,882 firm-year observations. The mean and standard deviation of JS distance are 0.27 and 0.14 respectively. The minimum and maximum of JS distance are 0.003 and 0.625. Table 2.5 on page 71 presents the correlations between other measures that are often used as a proxy for firm-level uncertainty in the literature (see Garfinkel 2009 and the references therein). I compare the firm-level JS distance against the four other firm-level measures: firm age, volatility of analyst's median forecast, and time-averaged difference between the lowest and highest analyst's forecasts. The JS distance is positively correlated (0.033) with the time-averaged difference between the lowest and highest analyst's forecasts for the firm's EPS, and is negatively correlated (-0.167) with the firm's age. The estimated coefficients of correlation are statistically significant at the 5% level. The estimate of correlation coefficient between the JS distance and the volatility of median forecasts is negative (-0.002), but statistically insignificant at the 10% level. Note that the earnings coefficient response (ERC) is also often used for measuring a firm's time-varying amount of uncertainty (Pastor et al., 2009). However, I run between-group regressions of leverage, where all the dependent and

independent variables are averaged over all years in the sample period. I find that time-averaged ERC is highly correlated with the time-averaged difference between the lowest and highest analyst's forecasts.

Table 2.6 on page 72 presents the pairwise correlations between the firm-characteristic variables and JS distance. JS distance is negatively correlated with leverage, profitability, tangibility, and firm size, while positively correlated with the firm's cash holding ratio and market-to-book ratio. These correlations are intuitive. As the amount of uncertainty of a firm increases, the firm tends to take on less leverage. Young and small firms are likely more uncertain than mature firms, due to the lack of historical observations of their profitability. I argue that cash is less risky, as well as a less uncertain asset than debt. Hence, when the firm faces higher uncertainty, its cash holdings more likely increase. A high market-to-book ratio implies high growth opportunities. These opportunities likely result from a large investment in innovative technology. Hence, the positive relation between a firm's market-to-book ratio and its uncertainty seems reasonable.

## 2.3.3 Between-firm leverage regression with new measure of uncertainty

In order to investigate the relationship between a firm's uncertainty as measured by the JS distance and its leverage, I run the following between-firm regression of leverage on the firm-level JS distance with industry fixed effects:

$$\overline{Leverage}_i = \alpha + \beta \cdot JSdist_i + \gamma'\overline{X}_i + \lambda \cdot vol(Profit.)_i + \delta \cdot \eta_j + \epsilon_i, \qquad (2.4)$$

where $\overline{Leverage}_i$ is firm $i$'s leverage averaged over all years in the sample period (1971-2004). The time-averaged firm-specific control variables $\overline{X}_i$ include profitability, tangibility, market-to-book ratio, and firm size. $vol(Profit.)_i$ is firm $i$'s volatility of profitability during the sample period, and $\eta_j$ is the industry-specific mean leverage, where $j$ indexes a specific industry in the 48 Fama-French industries. $JSdist_i$ is the amount of uncertainty faced by firm $i$.

To obtain the JS distance at the firm level (i.e., the amount of uncertainty faced by the firm), I first collect the quarterly earnings of each firm from 1970 to 2008. Then, I estimate the mixture density and the single normal density fitted to these samples. I compute the JS distance between the estimated mixture and single density. To ensure the convergence of EM algorithms, I only include the firms with at least 12 valid realized earnings during the

sample period. I also winsorize the JS distance variable at the top and bottom 1%.

Table 2.7 on page 73 provides the results of regression estimates of (2.4) on the preceding page. Models 1 and 2 present the benchmark results without the firm's JS distance. The signs of the estimates are as expected. A firm with high market-to-book ratio and high profitability takes on low leverage, whereas a firm with large sales and high tangible assets takes on high leverage. In Model 2 for market leverage, the volatility of profitability is negatively related to market leverage although the estimate is not statistically significant at the 10% level. Firms with more volatile assets will have higher default probabilities and expected costs of bankruptcy. Hence, they are expected to choose low leverage. However, it is counterintuitive that in Model 1 for book leverage, the estimate of the volatility of profitability is positive and it is statistically significant at the 10% level.

Models 3 and 4 in Table 2.7 present the results of between-group leverage regressions with JS distance in (2.4). In Model 3, book leverage is the dependent variable; market leverage is the dependent variable in Model 4. The coefficient estimates of JS distance in Models 3 and 4 enter with negative signs. The $t$-statistics (-3.4 in Model 3 and -4.1 in Model 4) provide evidence that they are statistically significant at the 1% level. High uncertainty measured by JS distance is associated with low leverage. In addition, the signs of coefficient estimates of the other control variables are consistent with intuitions, as in the benchmark models (Models 1 and 2).

To assess the economic magnitude of the effect of uncertainty on the firm's mean leverage, consider a median firm in the sample whose JS distance is 0.253. Suppose this firm's JS distance increases by one standard deviation (0.14). Then, its mean market leverage decreases by approximately $0.14 \times 0.051$ or 0.71%. Scaled by the mean market leverage in the sample (0.2), a one standard deviation increase of the firm's JS distance decreases its mean market leverage by approximately 3.6%. The same procedure leads to about a 3% fall in mean book leverage.

The coefficient estimate of volatility of profitability in Model 4 in Table 2.7 is statistically insignificant at the 10% level. In Model 3, the coefficient of the volatility of profitability is statistically significant but enters with a positive sign. Whereas, in both Models 3 and 4, uncertainty (i.e., JS distance) is negatively associated with leverage. The coefficient estimates of JS distance are statistically significant at the 1% level. The results suggest that uncertainty provides a plausible explanation for a firm's taking on low leverage.

It is not certain whether or not there is severe correlation between the JS distance and the volatility of profitability. The high correlation between covariates causes the instability

of the estimates, as well as the inefficiency of the estimates. I find no statistical evidence of severe multicollinearity between the JS distance and the volatility of profitability. In Model 3, the variance inflation factors (VIF) of all covariates excluding the industry dummies are reported from 1.11 to 3.94, which are under the threshold of severe multicollinearity (Greene, 2012, pg. 130).[1] Model 4 reports similar VIF results.

## 2.3.4   Additional test: uncertainty and cash holding

I argue that cash is less risky, as well as a less uncertain asset, than debt. Hence, the prediction is that the high uncertainty of a firm is likely associated with an increase in the firm's cash holdings. To validate this prediction, I run the between-firm regression of a firms's cash holding ratio on the firm's JS distance, along with the firm-specific control variables including the volatility of a firm's profitability. Table 2.8 on page 74 presents the results. In Model 2, the coefficient estimate of the JS distance is positive and statistically significant at the 1% level. Consider a median firm whose cash holding ratio is 0.26. A one standard deviation increase in JS distance is associated with an increase in the cash holding ratio by 1.23% (0.14 × 0.088). Scaling by the cash holding ratio (0.26), a one standard deviation increase in uncertainty is associated with an approximately 4.7% increase in its cash holding ratio. The result is economically sizeable.

## 2.3.5   Uncertainty and almost zero leverage firms

This section shows that the propensity of a firm's taking almost zero leverage is positively correlated with the firm's JS distance. Following Strebulaev and Yang (2012), I classify a firm as an almost zero leverage-firm (AZL) if its time-averaged leverage is less than 5%. Table 2.9 on page 75 presents the percentages of the almost zero market leverage firms in the 48 Fama-French industries. For instance, there are 661 firms in the Drugs industry of the 48 Fama-French industries during the sample period 1970-2008; 309 (46.7%) firms take on less than or equal to 5% time-averaged market leverage. There are 11,611 firms in the sample; approximately 20% of them are almost zero market leverage firms.

   The model used to estimate the relationship between the JS distance and the propensity

---

[1]Though VIF is an imperfect measure of the multicollinearity between covariates, it can exclude severe multicollinearity among the covariates. By a rule of thumb, multicollinearity is severe if VIF is higher than 10.

of firms' taking almost zero leverage is:

$$\text{logit}(\mathbb{P}[AZL_i|\Xi_i]) = \alpha + \beta \cdot JSdist_i + \gamma'\overline{X}_i + \lambda \cdot vol(Prof.)_i + \eta_j, \qquad (2.5)$$

where $AZL_i$ is a dummy variable that is 1 if a firm has less than 5% time-averaged leverage, and $\Xi_i$ is a vector of the independent variables: $\Xi_i = (JSdist_i, Vol(Prof.)_i, \overline{X}_i, \eta_j)$. The time-averaged firm-specific control variables $\overline{X}_i$ are the same as in the model (2.4) on page 58: profitability, tangibility, market-to-book ratio, and firm size. $vol(\text{Profit.})_i$ is firm $i$'s volatility of profitability during the sample period and $\eta_j$ is the industry fixed-effects, where $j$ indexes a specific industry in the 48 Fama-French industries. $JSdist_i$ is the amount of uncertainty faced by an individual firm $i$.

Table 2.10 on page 76 presents the results of logit regressions, where the dependent binary variable is 1 if a firm's time-averaged market leverage is less than 5%: in Models 1 and 3, the variable AZ-BL is 1 if a firm's time-averaged book leverage is less than 5%; in Models 2 and 4, the variable AZ-ML is 1 if a firm's time-averaged market leverage is less than 5%. The regression results suggest that the firms tend to take almost zero leverage, when they are small, to have high market-to-book ratios and profitability, and to have less tangible assets. Based on the sufficiently high $t$-statistics, these relationships are statistically significant at the 1% level. When the firm faces higher uncertainty as measured by the JS distance, it is more likely to take on almost-zero leverage. The economic signicance of the JS distance is measured by the change in propensity of a firm's taking almost zero leverage for the change in its JS distance. Fixing the other covariates as their mean values, an increase in JS distance from one standard deviation below to one standard deviation above its mean is associated with an increase in the propensity of firms' taking on almost zero market (book) leverage by 2.4% (3.2%); an impact of changing JS distance from the minimum to the maximum is associated with an increase in the propensity of firms' taking an almost zero market (book) leverage policy by 6.3% (7.2%). The regression estimates of JS distance are also statistically significant at the 1% level. It is counter-intuitive to observe the coefficient estimates of the volatility of profitability that enter with negative signs in Models 3 and 4. However, Strebulaev and Yang (2012) also find the similar empirical results (see Table 3 in Strebulaev and Yang 2012).

## 2.4 Earnings Response Coefficient as a Proxy for Uncertainty

In this section, instead of using the new measure of uncertainty, I use the earnings response coefficient (ERC) – the stock price reaction to an earnings announcement– as an empirical proxy for a firm's uncertainty. The rationale is as follows. Because they use Bayesian updating, investors who are most uncertain about the profitability of a firm (investors' precision of prior is very low) should respond strongly to earnings surprises (*i.e.,* update their posterior belief by more) (Pastor et al., 2009). In the literature, ERC is widely used as a proxy for the uncertainty of the firm's future profitability (Pastor et al., 2009; Cremers and Yan, 2009; He et al., 2010). However, as I mentioned, my new measure is superior to the ERC because the ERC is a proxy, not a direct measure, for the amount of uncertainty that the firm faces. However, the purpose of this section is to confirm that there is a negative relationship between uncertainty when I use the alternative proxy for the uncertainty.

I obtain firm-quarter observations from the Compustat quarterly database of publicly traded US firms from 1990 to 2009. I again exclude the utility and financial firms. To construct ERC, I follow the details in Pastor et al. (2009). They use two different ERC measures: ERC(1) is the average of the firm's previous 12 stock price reactions to quarterly earnings surprises and ERC(2) is the negative of the regression slope of the firm's last 20 quarterly earnings surprises on its abnormal stock returns on the firm's earnings announcements. In this paper, I use ERC(1) as a proxy for the manager's ambiguity about the firm's prospects.

Specifically, I obtain RC, which is the abnormal return due to a quarterly earnings announcement divided by the unexpected quarterly earnings. The abnormal return is measured as the cumulative return of stock $i$ in excess of the industry return starting one trading day before the firm's earnings announcement and ending one trading day after the same announcement. Quarterly earnings announcement dates are obtained from the IBES dataset. The industry returns are the daily returns of the 48 value-weighted industry portfolios available from Ken French's website. The unexpected quarterly earnings are equal to the difference between the actual quarterly earnings per share (obtained from the IBES unadjusted actuals file) and the mean of all analyst forecasts of EPS using the latest forecasts in IBES for the given fiscal quarter, deflated by the book equity per share of the firm. Because RCs are noisy, I require at least six valid observations of RCs. I trim RC at 5% and 95%, and average the trimmed quarterly RCs over the rolling three-year window to obtain ERC(1).

As stated by Pastor et al. (2009), earnings surprises and stock returns should have the same sign, that is, ERC(1) should be positive. Consistent with their finding, I also find

that 24% of my sample are of negative ERC(1). Hence, I construct from ERC(1) two new variables: (i) $ERC^+$ if ERC(1) > 0 otherwise zero and (ii) $ERC^-$ if ERC(1) < 0 otherwise zero. Table 2.11 provides the summary statistics of ERC(1) in my sample. The mean is 4.33, the median 3.43, the standard deviation 5.98, the minimum is -14.65, and the maximum is 30.64. I take nonlinear (log) transformation of $ERC(1)$ to balance its scale to that of the dependent variable, market (or book) leverage.

I adopt a firm fixed effect regression, which enables me to investigate the relationship between the firm's uncertainty and its leverage in a given time. I estimate the following model:

$$Leverage_{it} = \alpha + \beta^+ \ln(ERC_{it}^+) + \beta^- \ln(ERC_{it}^-) + \gamma X_{it} + \eta_i + \nu_t + \epsilon_{it}, \qquad (2.6)$$

where $Leverage_{it}$ is the book or market leverage of firm $i$ at time $t$, $\eta_i$ is a firm fixed effect, $\nu_t$ a year fixed effect, $\ln(ERC_{it}^+)$ and $\ln(ERC_{it}^-)$ the uncertainty about the firm's earnings at time $t$, and $X_{it}$ a set of firm level controls measured at time $t$, and $\epsilon_{it}$ an error term.

The firm-specific control variables are the profitability of the firm at time $t$, the market-to-book ratio at time $t$, the tangibility at time $t$, the median leverage in the industry at time $t$, the log of size at time $t$, and the volatility of firm's past three-years earnings. Table 2.1 provides the variable definitions.

Table 2.12 provides the results of estimated model (2.6). All standard errors are clustered at the firm level. Models 1 and 2 represent the benchmark results with the control variables and firm fixed effects but without the firm's uncertainty measures ($ERC^+$ and $ERC^-$). All the signs and the order of magnitudes of the coefficient estimates are consistent with the previous literature (for example, Lemmon et al. (2008)). That is, a firm with high market-to-book ratio and high profitability has lower book and market leverage, whereas a firm with large sales revenue has higher book and market leverage.

Models 3 and 4 present the results of firm-fixed effect panel regressions with ERC. In Model 3, book leverage is the dependent variable and market leverage is the dependent variable in Model 4. The coefficients of the ERC variables ($\ln(ERC^+)$ and $\ln(ERC^-)$) in Models 3 and 4 enter with a negative sign. The $t$-statistics (-5.8 in Model 3 and -3.4 in Model 4) provide evidence that they are statistically significantly different from zero. That is, firms experiencing high uncertainty reduce leverage. In addition, all estimates of the other control variables are of the expected sign and statistically significant.

To assess the economic magnitude of the effect of uncertainty on the firm's leverage, consider a median firm in the sample whose $\ln(ERC^+)$ is 1.842. Suppose this firm's uncertainty

increases in a given year by one standard deviation (0.8). Then, its book leverage decreases by approximately $0.8 \times 0.006$ or approximately 0.5%. Scaling by the average book leverage in the sample 0.21, I find that one standard deviation increase of the firm's uncertainty measured by $\ln(ERC^+)$ leads to a 2.4% fall in book leverage. The same procedure leads to a 2.3% fall in market leverage. These results are sizable.

Notice that the estimated coefficient of volatility of EBIT in Models 3 is reported as close to zero and statistically insignificant ($t$-statistic 0.8). In Model 4, the estimated coefficient of volatility of EBIT is statistically significant but enters with a positive sign. On the other hand, uncertainty has a statistically significant negative effect on firm's leverage in Models 3 and 4. Although multicollinearity between the volatility and uncertainty is possible, I find no statistical evidence of severe multicollinearity between ERC and volatility of EBIT. In Model 3, the variance inflation factors (VIF) of all covariates are reported from 1.02 to 1.77, and the mean VIF of the model is reported as around 1.73, which excludes the possibility of severe multicollinearity between the covariates. Model 4 reports similar VIF results.

## 2.5   Conclusion

I consider a manager who faces Knightian uncertainty about the firm;s profitability in the next period. I develop a new empirical measure for the amount of Knightian uncertainty faced by the manager. I propose that a mixture of normal distributions characterizes the case where the manager is unsure about the distribution of the firm's future profitability, whereas a single normal distribution characterizes the case when she is absolutely sure about the distribution. I measure the Jensen-Shannon distance (JS distance) between the single normal distribution and the mixture of three normal distributions fitted to the firm's realized profitability samples. This JS distance is the amount of uncertainty faced by the firm.

Using this new measure, I run between-firm regressions of leverage on uncertainty along with other firm-specific control variables including the volatility of profitability. When the JS distance increases by one standard deviation, the median firm in the sample increases its market leverage by approximately 3.6%. Finally, I run logit regressions of the propensity of a firm having almost zero leverage on its JS distance. The results suggest that changing the JS distance from the minimum to the maximum is associated with an increase in the propensity of a firm having almost zero market (book) leverage by 6.3% (7.2%). Therefore, the amount of uncertainty perceived by a firm is an important determinant of its leverage.

## 2.A  Jensen-Shannon Statistical Distance

The complete discussion of various statistical distances can be found in Cover and Thomas (2006). The relative entropy or the Kullback-Leibler divergence between to two probability functions $p(x)$ and $q(x)$ is defined by

$$D(p||q) = \int p(x) \log \frac{p(x)}{q(x)} dx \tag{2.7}$$

The KL divergence appears in statistics as the expected value of the log likelihood ratio. It therefore determines the ability to discriminate between two distributions $p(x)$ and $q(x)$. However, this is a quasimetric due to the asymmetric nature. Hence, we may use the symmetrized version of KL divergence known as the Jensen-Shannon divergence. That is

$$JSD(p||q) = \pi_1 D(p||m) + \pi_2 D(q||m) \tag{2.8}$$

where $\pi_1 + \pi_2 = 1$ and $m = \pi_1 p + \pi_2 q$. For the complete symmetry, I set $\pi_1 = \pi_2 = \frac{1}{2}$. Finally, the Jensen-Shannon metric is defined as the square root of Jensen-Shannon divergence.

## 2.B  Expectation Maximization Algorithms

A dataset composed of $N$ datapoints is considered. For a mixture of $K$ components of normal distributions, the probability of a datapoint $x_j$ is defined by

$$P(x_j) = \sum_{k=1}^{K} \pi_k P(x_j|k) = \sum_{k=1}^{K} \pi_k N(\mu_k, \sigma_k^2) \tag{2.9}$$

where $\pi_k \in [0,1]$ is a prior satisfying $\sum_{k=1}^{K} \pi_k = 1$. The mixture of normal distributions is then described by the set of parameters $\theta = \{\pi_k, \mu_k, \sigma_k\}_{k=1}^{K}$, which needs to be estimated from the observations $\{x_j\}_{j=1}^{N}$

The parameters of the mixture of normal distributions can be estimated via maximizing its log-likelihood function with observation $x = \{x_j\}_{j=1}^{N}$:

$$\mathcal{L}(\theta|x) = \sum_{j=1}^{N} \log P(x_j) \tag{2.10}$$

Conventionally, we can find ML estimates of $\theta$ using a direct method such as gradient-

based numerical optimization. Dempster et al. (1977) show that under fairly mild regularity conditions, the expectation-maximization (EM) algorithms converge to a local maximum of the mixture likelihood function.

First, we introduce hidden state variable $z_j$ ($1 \times K$ vector) associated with the observation $x_j$: $z_j$ is k-dimension $0/1$ vector with $(z_j)_i = z_{ij} = 1$ if $x_j$ arises from the $i$-th component of the mixture. Then, the log-likelihood function with this auxiliary variables $z_{ij}$ is

$$\mathcal{L}(\theta, z|x) = \sum_{j=1}^{N} \sum_{k=1}^{K} z_{kj} \left( \log \pi_k + \log N(\mu_k, \sigma_k) \right) \tag{2.11}$$

Dempster et al. (1977) show that EM algorithms are applied to this problem by treating $z_{ij}$ as missing (hidden) data. It proceeds iteratively in two steps, E (for expectation) and M (for maximization). The addition of the hidden variables to the problem is handled by the E-step. which takes the conditional expectation of the augmented log likelihood function (2.11), given the observed $x$, using the current fit for $\theta$. Then M-step on $u$-th iterations maximizes the augmented log-likelihood functions to provide the updated parameter $\theta^{(u+1)}$. More precisely,

E -step :

$$z_{kj}^u = \frac{\pi_k^u N(\mu_k^u, \sigma_k^u)}{\sum_{i=1}^{K} \pi_i^u N(\mu_i^u, \sigma_i^u)}$$

$$E_k^u = \sum_{j=1}^{K} z_{kj}^u$$

M -step :

$$\pi_k^{u+1} = \frac{E_k^u}{N}$$

$$\mu_k^{u+1} = \frac{\sum_{j=1}^{N} z_{kj}^u x_j}{E_k^u}$$

$$\sigma_k^{u+1} = \sqrt{\frac{\sum_{j=1}^{N} z_{kj}^u (x_j - \mu_{k+1})^2}{E_k^u}}$$

The iteration stops when the increase in the log-likelihood function at each iteration becomes too small. That is, $\frac{\mathcal{L}^{u+1}}{\mathcal{L}^u} < \epsilon$, for example $\epsilon = 10^{-4}$. The statistical literature asserts that the EM is more reliable than the direct ML estimation (McLachlan and Krishnan, 1997).

**Table 2.1 Variable Definitions**

| Variables | Definition |
|---|---|
| Book Equity | Stockholders Equity + Deferred Taxes and Investment Tax Credit - Preferred Stock (PSTKQ) if (PSTKQ) missing then Preferred Stock Redemption Value (PSTKRQ) |
| Total Debt | Long term debt + Short term debt |
| Market Equity | Common Shares Outstanding * Closing Price |
| Market Leverage | Total Debt / Market Value of Asset |
| Book Leverage | Total Debt / Total Asset |
| Market-to-Book | (Market Equity + Total Debt + Preferred stock value - Deferred taxes and investment tax credits) / Total Asset |
| Firm Size | log(Sales[millions of dollars(henceforth, MM$)]) |
| Tangibility | Plant Property and Equipment / Total Asset |
| Profitability | (Earnings Before Interest, Taxes, Depreciation and Amortization) / Total Asset |
| Vol(Prof.) | Volatility of profitability (minimum 5 years of data required) |

**Table 2.2 Summary Statistics of Firm-Quarter Observations, 1990-2009**

This table presents summary statistics of firm-quarter observations from Compustat quarterly database covering the period 1990-2009. The total the number of firms is 6,081. The panel is unbalanced. Profitability is EBIT/Asset, Tangibility PPE/Assets. The volatility of profitability is computed from the firm's past 12-quarters profitability observations.

| Variable | Mean | Median | Std. Dev. | Number of obs. |
|---|---|---|---|---|
| Market Leverage | 0.196 | 0.141 | 0.199 | 109131 |
| Book Leverage | 0.207 | 0.192 | 0.176 | 109131 |
| Market-to-Book | 1.671 | 1.256 | 1.315 | 99031 |
| Profitability | 0.028 | 0.033 | 0.042 | 109131 |
| Vol. of Profitability | 0.34 | 0.275 | 1.279 | 108958 |
| Tangibility | 0.295 | 0.231 | 0.228 | 109032 |
| Log(firm age[years]) | 2.572 | 2.546 | 0.843 | 109131 |
| Log(R&D[MM$]) | 0.025 | 0.016 | 0.034 | 55657 |
| Log(Sales[MM$]) | 4.839 | 4.833 | 1.847 | 108629 |

**Table 2.3 Summary Statistics of Firm-level Jensen-Shannon Distance**

This table represents the summary statistics of the firm-level Jensen-Shannon Distance (the JS distance). Only firms with at least 12 valid quarterly earnings observations during 1970-2008 are included. There are 11,882 firms in the filtered sample.

| | JS Distance |
|---|---|
| N | 11882 |
| mean | 0.270 |
| sd | 0.139 |
| min | 0.003 |
| p5 | 0.071 |
| p25 | 0.164 |
| p50 | 0.253 |
| p75 | 0.363 |
| p95 | 0.526 |
| max | 0.625 |

Table 2.4: **Jensen-Shannon Distances in the 48 Fama-French Industries**
This table represents (i) the parameter estimates of the mixture of three normal distributions, (ii) the parameter estimates of the single normal distribution, and (iii) the Jensen-Shannon distance (the JS distance) between the mixture of normal distributions and the single distribution in the 48 Fama-French industries. $\{(\theta_i, \sigma_i)\}_{i=1}^3$ are the parameter estimates of $i$-th component of the mixture and $\{\mu_i\}_{i=1}^3$ is the manager's belief attached to each $i$-th component of the mixture. $(\theta_0, \sigma_0)$ are the parameter estimates of the single normal distribution representing the case where the manager faces no uncertainty. The JS distance represents the statistical distance between the mixture of normal distributions and the single normal distribution. The rows of the table are sorted in descending order of the JS distances.

| | Parameters of Mixture Components | | | | | | | | | Single Normal | | |
| | Means | | | Variances | | | Belief | | | | | |
| 48-FF | $\theta_1$ | $\theta_2$ | $\theta_3$ | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ | $\mu_1$ | $\mu_2$ | $\mu_3$ | $\theta_0$ | $\sigma_0$ | JSdist |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mines | -0.021 | 0.033 | -0.199 | 0.042 | 0.026 | 0.130 | 0.351 | 0.515 | 0.134 | -0.017 | 0.095 | 0.312 |
| Gold | -0.237 | 0.003 | -0.073 | 0.124 | 0.029 | 0.052 | 0.101 | 0.683 | 0.217 | -0.037 | 0.090 | 0.299 |
| MedEq | -0.171 | 0.037 | -0.039 | 0.089 | 0.022 | 0.047 | 0.199 | 0.514 | 0.288 | -0.026 | 0.093 | 0.286 |
| Chems | -0.043 | -0.188 | 0.039 | 0.056 | 0.025 | 0.021 | 0.148 | 0.014 | 0.837 | 0.023 | 0.048 | 0.279 |
| BusSv | -0.044 | -0.192 | 0.034 | 0.050 | 0.077 | 0.028 | 0.229 | 0.077 | 0.694 | -0.001 | 0.075 | 0.265 |
| Comps | -0.157 | -0.033 | 0.035 | 0.069 | 0.045 | 0.026 | 0.113 | 0.248 | 0.640 | -0.003 | 0.073 | 0.260 |
| ElcEq | 0.037 | -0.027 | -0.131 | 0.021 | 0.038 | 0.033 | 0.755 | 0.201 | 0.044 | 0.017 | 0.048 | 0.250 |
| LabEq | 0.035 | -0.139 | -0.029 | 0.022 | 0.029 | 0.040 | 0.754 | 0.039 | 0.207 | 0.015 | 0.048 | 0.234 |
| Soda | -0.031 | 0.076 | 0.033 | 0.057 | 0.030 | 0.013 | 0.242 | 0.211 | 0.547 | 0.026 | 0.049 | 0.230 |
| Toys | -0.218 | 0.031 | -0.067 | 0.036 | 0.033 | 0.061 | 0.027 | 0.807 | 0.166 | 0.008 | 0.065 | 0.227 |
| Drugs | -0.066 | -0.248 | 0.039 | 0.062 | 0.131 | 0.021 | 0.547 | 0.214 | 0.239 | -0.080 | 0.124 | 0.226 |
| Mach | -0.028 | 0.033 | -0.132 | 0.039 | 0.022 | 0.031 | 0.147 | 0.822 | 0.031 | 0.019 | 0.043 | 0.224 |
| Chips | -0.138 | 0.034 | -0.026 | 0.034 | 0.024 | 0.040 | 0.041 | 0.737 | 0.222 | 0.013 | 0.049 | 0.210 |
| Telcm | -0.002 | 0.042 | -0.103 | 0.031 | 0.020 | 0.037 | 0.293 | 0.650 | 0.057 | 0.021 | 0.044 | 0.210 |
| Smoke | 0.043 | 0.170 | 0.057 | 0.021 | 0.025 | 0.065 | 0.590 | 0.088 | 0.322 | 0.058 | 0.054 | 0.208 |
| Hlth | -0.100 | 0.038 | -0.016 | 0.021 | 0.021 | 0.035 | 0.024 | 0.814 | 0.162 | 0.026 | 0.036 | 0.197 |
| Autos | 0.080 | -0.024 | 0.034 | 0.011 | 0.038 | 0.019 | 0.050 | 0.113 | 0.836 | 0.030 | 0.030 | 0.150 |
| Hshld | 0.037 | 0.037 | -0.032 | 0.025 | 0.025 | 0.039 | 0.437 | 0.414 | 0.149 | 0.027 | 0.037 | 0.147 |
| Guns | 0.043 | -0.018 | 0.028 | 0.027 | 0.046 | 0.018 | 0.202 | 0.119 | 0.679 | 0.025 | 0.030 | 0.143 |
| Oil | 0.037 | -0.102 | -0.007 | 0.024 | 0.023 | 0.034 | 0.699 | 0.027 | 0.274 | 0.021 | 0.039 | 0.140 |
| Whlsl | -0.037 | 0.027 | 0.026 | 0.028 | 0.020 | 0.021 | 0.099 | 0.893 | 0.008 | 0.020 | 0.028 | 0.140 |
| Aero | 0.067 | -0.020 | 0.031 | 0.010 | 0.024 | 0.014 | 0.086 | 0.088 | 0.826 | 0.030 | 0.023 | 0.140 |
| Rubbr | -0.020 | 0.037 | -0.080 | 0.029 | 0.021 | 0.010 | 0.098 | 0.891 | 0.011 | 0.030 | 0.029 | 0.139 |
| Beer | 0.042 | 0.011 | -0.056 | 0.022 | 0.028 | 0.012 | 0.786 | 0.182 | 0.032 | 0.033 | 0.030 | 0.122 |
| Fun | 0.106 | 0.030 | -0.022 | 0.018 | 0.027 | 0.053 | 0.054 | 0.673 | 0.273 | 0.020 | 0.047 | 0.111 |
| PerSv | -0.054 | 0.033 | -0.097 | 0.024 | 0.029 | 0.011 | 0.047 | 0.942 | 0.011 | 0.027 | 0.036 | 0.106 |

Continued on next page

Table 2.4 **Jensen-Shannon Distances in the 48 Fama-French Industries** (Continued)

| 48-FF | Means | | | Variances | | | Belief | | | Single Normal | | JSdist |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\theta_1$ | $\theta_2$ | $\theta_3$ | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ | $\mu_1$ | $\mu_2$ | $\mu_3$ | $\theta_0$ | $\sigma_0$ | |
| Boxes | 0.036 | 0.068 | -0.015 | 0.014 | 0.015 | 0.017 | 0.850 | 0.104 | 0.046 | 0.037 | 0.020 | 0.100 |
| Books | -0.033 | 0.039 | 0.093 | 0.021 | 0.023 | 0.015 | 0.081 | 0.838 | 0.080 | 0.037 | 0.034 | 0.099 |
| Coal | 0.024 | 0.081 | -0.035 | 0.020 | 0.008 | 0.011 | 0.919 | 0.026 | 0.055 | 0.022 | 0.025 | 0.091 |
| Meals | -0.038 | 0.040 | 0.014 | 0.010 | 0.020 | 0.025 | 0.025 | 0.713 | 0.262 | 0.032 | 0.026 | 0.086 |
| Trans | 0.034 | 0.035 | -0.024 | 0.021 | 0.021 | 0.018 | 0.746 | 0.186 | 0.069 | 0.030 | 0.025 | 0.086 |
| Food | -0.021 | 0.081 | 0.035 | 0.019 | 0.013 | 0.020 | 0.076 | 0.079 | 0.845 | 0.034 | 0.028 | 0.077 |
| Agric | 0.017 | 0.073 | -0.049 | 0.025 | 0.025 | 0.027 | 0.705 | 0.119 | 0.176 | 0.012 | 0.042 | 0.075 |
| Cnstr | 0.021 | 0.021 | -0.047 | 0.023 | 0.023 | 0.015 | 0.283 | 0.672 | 0.045 | 0.018 | 0.027 | 0.074 |
| BldMt | 0.036 | 0.042 | -0.033 | 0.024 | 0.026 | 0.014 | 0.960 | 0.000 | 0.039 | 0.034 | 0.027 | 0.068 |
| Rtail | 0.031 | -0.032 | 0.089 | 0.024 | 0.016 | 0.012 | 0.886 | 0.059 | 0.055 | 0.031 | 0.030 | 0.064 |
| Clths | 0.035 | -0.030 | 0.105 | 0.027 | 0.014 | 0.012 | 0.907 | 0.054 | 0.040 | 0.034 | 0.033 | 0.059 |
| Ships | 0.080 | -0.012 | 0.031 | 0.010 | 0.011 | 0.018 | 0.035 | 0.067 | 0.898 | 0.030 | 0.022 | 0.055 |
| FabPr | 0.069 | 0.032 | -0.011 | 0.012 | 0.018 | 0.011 | 0.074 | 0.855 | 0.070 | 0.032 | 0.022 | 0.047 |
| Txtls | 0.077 | -0.005 | 0.034 | 0.013 | 0.010 | 0.018 | 0.049 | 0.057 | 0.894 | 0.034 | 0.022 | 0.042 |
| Steel | 0.029 | 0.029 | -0.023 | 0.022 | 0.022 | 0.010 | 0.261 | 0.701 | 0.038 | 0.027 | 0.023 | 0.041 |
| Paper | 0.038 | 0.038 | 0.038 | 0.022 | 0.022 | 0.022 | 0.187 | 0.647 | 0.166 | 0.038 | 0.021 | 0.011 |

**Table 2.5 Pairwise Correlations between the Other Measures for proxy for Uncertainty and Jensen-Shannon Distance**
This table presents correlations between other measures as a proxy for firm-level uncertainty. Comparing the JS distance against four other measures: firm's age, volatility of analyst's median forecast (Vol(Med. Forecast)), and time-averaged difference between the lowest and highest analyst's forecasts for a firm's earnings per share (High - Low). The number in parenthesis represents $p$-value of each correlation coefficient.

| Variables | JS Distance | vol(Prof.) | High-Low | vol(Med. Forecast) | Firm Age |
|---|---|---|---|---|---|
| JS Distance | 1.000 | | | | |
| vol(Prof.) | 0.271 | 1.000 | | | |
| | (0.000) | | | | |
| High-Low | 0.033 | 0.083 | 1.000 | | |
| | (0.009) | (0.000) | | | |
| vol(Med. Forecast) | -0.002 | 0.149 | 0.731 | 1.000 | |
| | (0.899) | (0.000) | (0.000) | | |
| Firm Age | -0.167 | -0.111 | 0.036 | 0.117 | 1.000 |
| | (0.000) | (0.000) | (0.005) | (0.000) | |

**Table 2.6 Pairwise Correlations between Firm-characteristic Variables and Jensen-Shannon Distance**

This table presents the pairwise correlations between the firm-characteristic variables and the Jensen-Shannon distance (the JS distance). The definitions of variables are detailed in Table 2.1. The number in parenthesis represents $p$-value of each correlation coefficient.

| Variables | BL | ML | Cash/Asset | Profitility | Mkt-to-Bk | Tangibility | Firm Size | vol(Prof.) | JS Dist. |
|---|---|---|---|---|---|---|---|---|---|
| Book Lev. | 1.000 | | | | | | | | |
| Market Lev. | 0.734 | 1.000 | | | | | | | |
| | (0.000) | | | | | | | | |
| Cash/Asset | -0.412 | -0.548 | 1.000 | | | | | | |
| | (0.000) | (0.000) | | | | | | | |
| Profitability | -0.227 | 0.171 | -0.304 | 1.000 | | | | | |
| | (0.000) | (0.000) | (0.000) | | | | | | |
| Mkt-to-Bk | 0.181 | -0.284 | 0.297 | -0.814 | 1.000 | | | | |
| | (0.000) | (0.000) | (0.000) | (0.000) | | | | | |
| Tangibility | 0.283 | 0.329 | -0.417 | 0.181 | -0.159 | 1.000 | | | |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | | | | |
| Firm Size | -0.057 | 0.229 | -0.342 | 0.613 | -0.507 | 0.115 | 1.000 | | |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | | | |
| vol(Prof.) | 0.174 | -0.195 | 0.254 | -0.759 | 0.694 | -0.168 | -0.557 | 1.000 | |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | | |
| JS Dist. | 0.004 | -0.098 | 0.154 | -0.189 | 0.156 | -0.075 | -0.225 | 0.271 | 1.000 |
| | (0.665) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | |

**Table 2.7 Jensen-Shannon Distance and Leverage**

This table presents the results of the between-group regression of leverage on the JS distance (i.e., the amount of uncertainty faced by the firm), the volatility of the firm's profitability, and firm-specific control variables with industry median leverage. Industry is classified using the 48 Fama-French classification. The definitions of variables are detailed in Table 2.1 on page 67. Standard errors are computed being robust to heteroscedasticity. The t-statistics of the coefficient estimates are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by "*", "**", and "***", respectively.

| | Model 1 Book Lev. | Model 2 Market Lev. | Model 3 Book Lev. | Model 4 Market Lev. |
|---|---|---|---|---|
| JS Distance | | | -0.044** | -0.051*** |
| *(Amount of Uncertainty)* | | | (-2.559) | (-3.073) |
| Profitability | -0.114*** | -0.119*** | -0.114*** | -0.119*** |
| | (-7.525) | (-5.345) | (-7.619) | (-5.394) |
| Market-to-Book | -0.000 | -0.022*** | -0.000 | -0.022*** |
| | (-0.132) | (-6.787) | (-0.199) | (-6.877) |
| Firm Size | 0.004 | 0.006** | 0.004 | 0.006** |
| | (1.409) | (2.228) | (1.412) | (2.257) |
| vol(Prof.) | 0.024* | -0.017 | 0.028** | -0.011 |
| | (1.986) | (-1.253) | (2.477) | (-0.834) |
| Tangibility | 0.223*** | 0.185*** | 0.222*** | 0.184*** |
| | (3.914) | (3.796) | (3.970) | (3.874) |
| Industry Mean BL | 0.844*** | | 0.844*** | |
| | (10.123) | | (10.142) | |
| Industry Mean ML | | 0.788*** | | 0.785*** |
| | | (15.892) | | (15.938) |
| Constant | -0.042* | 0.028* | -0.030 | 0.042*** |
| | (-1.900) | (1.900) | (-1.525) | (3.209) |
| R-squared | 0.211 | 0.282 | 0.212 | 0.283 |
| Num. of Obs. | 11566 | 11566 | 11566 | 11566 |

**Table 2.8 Additional Test: Jensen-Shannon Distance and Cash/Asset ratio**

This table presents the results of the between-group regression of Cash/Asset ratio on the Jensen-Shannon distance (the JS distance), the volatility of the firm's profitability, and firm specific controls with industry fixed effects. Industry is classified using the 48 Fama-French classification. Standard errors are computed being robust to heteroscedasticity. The t-statistics of the coefficient estimates are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by "*", "**", and "***", respectively.

|  | Model 1 Cash/Asset | Model 2 Cash/Asset |
|---|---|---|
| JS Distance |  | 0.088*** |
| *(Amount of Uncertainty)* |  | (8.426) |
| Profitability | -0.030*** | -0.030*** |
|  | (-5.528) | (-5.563) |
| Market-to-Book | 0.005*** | 0.005*** |
|  | (7.238) | (7.635) |
| Firm Size | -0.000 | -0.000 |
|  | (-0.218) | (-0.020) |
| vol(Prof.) | -0.009 | -0.018*** |
|  | (-1.430) | (-2.886) |
| Tangibility | -0.338*** | -0.333*** |
|  | (-38.281) | (-37.806) |
| Constant | 0.266*** | 0.237*** |
|  | (12.173) | (10.761) |
| Industry FE | Yes | Yes |
| R-squared | 0.411 | 0.414 |
| Num. of Obs. | 11566 | 11566 |

## Table 2.9 Almost Zero Market Leverage Firms

This table presents the percentage of almost zero market leverage firms in the 48 Fama-French industries. The table is sorted in descending order of the JS distance (i.e., the amount of uncertainty).

| FF 48 Industries | Almost Zero Lev. (%) | Num. of firms | JS Distance |
|---|---|---|---|
| Mines | 25.0% | 56 | 0.312 |
| Gold | 23.2% | 56 | 0.299 |
| MedEq | 41.1% | 438 | 0.286 |
| Chems | 15.2% | 211 | 0.279 |
| BusSv | 36.0% | 1713 | 0.265 |
| Comps | 34.4% | 607 | 0.26 |
| ElcEq | 12.4% | 186 | 0.25 |
| LabEq | 30.4% | 280 | 0.234 |
| Soda | 15.4% | 26 | 0.23 |
| Toys | 10.4% | 135 | 0.227 |
| Drugs | 46.7% | 661 | 0.226 |
| Mach | 12.7% | 434 | 0.224 |
| Chips | 27.7% | 689 | 0.21 |
| Telcm | 7.9% | 494 | 0.21 |
| Smoke | 0.0% | 7 | 0.208 |
| Hlth | 10.1% | 288 | 0.197 |
| Autos | 9.1% | 176 | 0.15 |
| Hshld | 12.6% | 239 | 0.147 |
| Guns | 20.0% | 20 | 0.143 |
| Oil | 11.6% | 704 | 0.14 |
| Aero | 8.8% | 57 | 0.14 |
| Whlsl | 8.8% | 554 | 0.14 |
| Rubbr | 8.5% | 153 | 0.139 |
| Beer | 10.0% | 40 | 0.122 |
| Fun | 9.1% | 286 | 0.111 |
| PerSv | 15.4% | 169 | 0.106 |
| Boxes | 2.6% | 39 | 0.1 |
| Books | 17.1% | 105 | 0.099 |
| Coal | 0.0% | 32 | 0.091 |
| Meals | 8.6% | 314 | 0.086 |
| Trans | 7.5% | 321 | 0.086 |
| Food | 8.5% | 211 | 0.077 |
| Agric | 12.2% | 49 | 0.075 |
| Cnstr | 4.1% | 170 | 0.074 |
| BldMt | 6.3% | 272 | 0.068 |
| Rtail | 11.7% | 737 | 0.064 |
| Clths | 15.6% | 179 | 0.059 |
| Ships | 7.7% | 26 | 0.055 |
| FabPr | 7.0% | 57 | 0.047 |
| Txtls | 2.9% | 105 | 0.042 |
| Steel | 2.4% | 170 | 0.041 |
| Paper | 5.5% | 145 | 0.011 |

**Table 2.10 Determinants of Almost Zero Leverage Policy**

This table presents the logit regression on the sample 1978-2006. The dependent variable is the dummy that equals 1 if the firm has leverage less than 5%: AZ-ML represents a dummy that equals 1 if the firm has market leverage less than 5%; AZ-BL represents a dummy that equals 1 if the firm has book leverage less than 5%. JS distance is the amount of uncertainty faced by the firm. All other independent variables are defined in Table 2.1 on page 67. Standard errors are computed being robust to heteroscedasticity. The t-statistics of the coefficient estimates are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by "*", "**", and "***", respectively.

|  | Model 1 AZ-BL | Model 2 AZ-ML | Model 3 AZ-BL | Model 4 AZ-ML |
|---|---|---|---|---|
| JS Distance | | | 0.878*** | 0.703*** |
| *(Amount of Uncertainty)* | | | (4.229) | (3.637) |
| Market-to-Book | 0.098*** | 0.225*** | 0.100*** | 0.227*** |
| | (8.136) | (18.958) | (8.245) | (19.079) |
| Profitability | 0.735*** | 1.067*** | 0.748*** | 1.070*** |
| | (6.754) | (11.258) | (6.833) | (11.287) |
| Firm Size | -0.063*** | -0.053*** | -0.065*** | -0.054*** |
| | (-3.649) | (-3.353) | (-3.730) | (-3.379) |
| vol(Prof.) | -0.973*** | -0.547*** | -1.060*** | -0.616*** |
| | (-7.001) | (-5.135) | (-7.510) | (-5.707) |
| Tangibility | -4.801*** | -3.859*** | -4.743*** | -3.833*** |
| | (-21.005) | (-20.713) | (-20.786) | (-20.612) |
| Industry Mean BL | -8.168*** | | -8.219*** | |
| | (-13.225) | | (-13.284) | |
| Industry Mean ML | | -7.094*** | | -7.082*** |
| | | (-21.623) | | (-21.567) |
| Constant | 1.474*** | 0.904*** | 1.254*** | 0.715*** |
| | (8.818) | (8.510) | (7.158) | (6.057) |
| R-squared | 0.137 | 0.187 | 0.139 | 0.188 |
| N | 11840 | 11840 | 11840 | 11840 |

**Table 2.11  Summary Statistics of Earnings Response Coefficient (ERC)**

When I compute the summary statistics of $\ln(ERC^+)$ and $\ln(ERC^-)$, I only include the entries that are non-zero. $\ln(ERC^+) = \ln(1 + ERC \cdot 1_{ERC>0})$ and $\ln(ERC^-) = -\ln(1 - ERC \cdot 1_{ERC<0})$.

| Stats | ERC | $\ln(ERC^+)$ | $\ln(ERC^-)$ |
|---|---|---|---|
| N | 109131 | 83177 | 25954 |
| Mean | 4.332 | 1.762 | -1.101 |
| Median | 3.427 | 1.842 | -1.109 |
| Std. Dev | 5.975 | 0.750 | 0.599 |
| Min | -14.646 | 0 | -2.750 |
| Max | 30.644 | 3.455 | 0 |

## Table 2.12 Earnings Response Coefficient (ERC) and Leverage

This table presents the results of panel regressions of leverage on the measures of uncertainty ($\ln(ERC^+)$ and $\ln(ERC^-)$) and the other firm specific control variables with firm fixed effects. $\ln(ERC^+) = \ln(1 + ERC \cdot 1_{ERC>0})$ and $\ln(ERC^-) = -\ln(1 - ERC \cdot 1_{ERC<0})$. $Profitability$ is the ratio $EBIT/Asset$, $Tangibility$ the ratio $PPENT/Assets$. Industry Median BL (ML) is the median book (market) leverage across all firms in the same industry at the same time. The parenthesis represents the $t$-statistic of the coefficient estimate. Standard errors are computed with clustering at firm level. The firm-quarter observations are taken from the Compustat quarterly database covering the period 1990 - 2009. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

|  | Model 1 BL | Model 2 ML | Model 3 BL | Model 4 ML |
|---|---|---|---|---|
| log(ERC+) |  |  | -0.006*** | -0.004*** |
| (**Amount of Uncertainty**) |  |  | (-5.576) | (-3.362) |
| log(ERC-) |  |  | 0.007*** | 0.005*** |
|  |  |  | (3.840) | (2.589) |
| vol(Prof.) | 0.000 | 0.001*** | 0.000 | 0.001*** |
|  | (0.378) | (2.963) | (0.316) | (2.929) |
| Market-to-Book | -0.010*** | -0.027*** | -0.010*** | -0.027*** |
|  | (-11.097) | (-23.266) | (-11.219) | (-23.306) |
| Profitability | -0.214*** | -0.408*** | -0.215*** | -0.409*** |
|  | (-9.317) | (-14.248) | (-9.370) | (-14.247) |
| Tangibility | 0.102*** | 0.104*** | 0.103*** | 0.105*** |
|  | (5.490) | (5.175) | (5.553) | (5.205) |
| Firm Size | 0.034*** | 0.039*** | 0.035*** | 0.040*** |
|  | (8.975) | (11.550) | (9.180) | (11.685) |
| Industry Med. BL | 0.377*** |  | 0.377*** |  |
|  | (15.857) |  | (15.896) |  |
| Industry Med. ML |  | 0.456*** |  | 0.456*** |
|  |  | (22.380) |  | (22.377) |
| Firm Fixed Effects | Yes | Yes | Yes | Yes |
| R-squared | 0.784 | 0.796 | 0.784 | 0.796 |
| N | 93192 | 93192 | 93192 | 93192 |

# Chapter 3

# Uncertainty, Real Options, and Dynamic Capital Structure

## 3.1 Introduction

Dixit (1992, page 190) describes the standard real options model as "a theory of optimal inertia" (or "a benevolent tyranny of the status quo"):

> " ...firms that refuse to invest even when the currently available rates of return are far in excess of the cost of capital may be optimally waiting to be surer that this state of affairs is not transitory. Likewise, farmers who carry large losses may be rationally keeping their operation alive on the chance that the future may be brighter."

Dixit and Pindyck (1994) show that an increase in the risk of the value of the project (in the sense of a mean preserving spread) raises the value of waiting and delays option exercise both in the case of irreversible investment and for an exit option.

However, Miao and Wang (2011) show that the standard real-option approach excludes the situation where agents are unsure about the likelihood of states of the world. Rather, the standard approach assumes that agents are absolutely certain about the stochastic process that determines the underlying state variable. They show that when an ambiguity-averse agent faces Knightian uncertainty about the evolution of the underlying project value, increased ambiguity lowers the value of waiting. Assuming that there is no uncertainty about the termination payoff after the option exercise, a more ambiguity-averse agent will exercise

the option earlier.[1]

To distinguish the effect of uncertainty from that of risk on exercising the option, I first provide two applications borrowed from Miao and Wang (2011): irreversible investment and firm exit. In Section 3.2, I extend their results from the discrete time to the continuous time setting. The continuous time formulation has distinct benefits over the discrete time counterpart. It allows for a closed form solution of the optimal option exercise boundary. Furthermore, the continuous time real options model with no uncertainty has important applications in the literature (see e.g., Leland (1994) on capital structure).

In Section 3.3, I incorporate Knightian uncertainty about the evolution of the firm's assets and investors' ambiguity-aversion into the model of Leland (1994). Here, I assume that ambiguity-averse investors display a max-min preference (Gilboa and Schmeidler, 1989), in contrast to the previous two chapters which assumed smooth ambiguity aversion (Klibanoff et al., 2005). I then show the ambiguity-aversion augmented model predicts substantially lower leverage than the benchmark model without uncertianty. Therefore, I propose that uncertainty provides a more plausible explanation for firms taking on low leverage than risk alone, even in a dynamic setting.

## 3.2  Risk, Knightian Uncertainty and Real Option

### 3.2.1  The set of priors in the dynamic setting[2]

Suppose $(X_t)$, the value of the underlying project, evolves as:

$$\frac{dX_t}{X_t} = \mu \, dt + \sigma \, dW_t, \tag{3.1}$$

where $(W_t)$ is a Brownian motion with respect to the reference probability measure $Q$, $\mu$ is the expected growth rate, and $\sigma$ is the volatility of the process. The manager is uncertain about the true measure $Q$. Instead, she considers a set of "perturbed" equivalent measures from the reference measure $Q$.[3]

---

[1]Miao and Wang (2011) also shows that if the agent is also uncertain about the termination payoffs, he may also exercise the option later. This is because ambiguity lowers both the termination payoff and the option value, and the decrease in the termination value dominates that in the option value.

[2]This section closely follows Epstein and Wang (1994) and Nishimura and Ozaki (2004).

[3]$Q_1$ and $Q_2$ are equivalent probability measures on $(E, \mathcal{F})$, if for any event A in $E$, $Q_1(A) = 0$ if and only if $Q_2(A) = 0$. Intuitively, this condition tells us that events under $Q_1$ are possible (or impossible) if and only if they are possible (or impossible) under $Q_1$.

Consider the probability space $(E, \mathcal{F}, Q)$ and the Brownian filtration $(\mathcal{F}_t)_{t \in [0,T]}$, where $\mathcal{F}_t$ is a the completed natural filtration of Brownian motion $(W_t)$ defined on $(E, \mathcal{F}, Q)$.[4] To construct the set of priors from the reference probability measure $Q$ on $(E, \mathcal{F})$, first define a $Q$-martingale process as follows. Consider a process $(\theta_t)$ that satisfies the Novikov condition. That is,

$$\mathbb{E}^Q \left[ \int_0^T \theta_s(\omega)^2 ds \right] < \infty.$$

Define

$$z_t^\theta := \exp\left( -\frac{1}{2} \int_0^t \theta_s^2 ds - \int_0^t \theta_s dW_s \right), \quad \text{such that } E^Q[Z_T] = z_0^\theta = 1.$$

In other words,

$$\frac{dz_t^\theta}{z_t} = -\theta_t dW_t.$$

Because $E^Q[z_T^\theta] = z_0^\theta = 1$ and $z_t^\theta > 0$ for all $t$, the Radon-Nykodym theorem (Duffie, 2001, pg. 332) ensures $(\theta_t)$ generates a probability measure $Q^\theta$ on $(E, \mathcal{F})$ that is equivalent to $Q$, where

$$dQ^\theta = z_t^\theta \cdot dQ \qquad \text{on } \mathcal{F}_t \text{ for all } t \in [0, T].$$

Thus, a set of priors $\Pi$ can be constructed by arbitrary selections of $\theta := (\theta_t)$

$$\Pi := \{Q^\theta\}.$$

Note that $(\theta_t) = 0$ represents no ambiguity (i.e., $\Pi$ is a singleton). By the Girsanov theorem (Karatzas and Shereve, 1988), we have a $Q^\theta$-Brownian motion $(W_t^\theta)$ such that

$$W_t^\theta = \int_0^t \theta_s \, ds + W_t.$$

When the manager is unsure about the true probability measure $Q$ on $(E, \mathcal{F})$, she considers the set of priors $\Pi$ and faces the set of dynamics

$$\frac{dX_t}{X_t} = (\mu - \sigma\theta_t)dt + \sigma dW_t^\theta. \tag{3.2}$$

When making a decision, the manager considers all stochastic differential equations (SDE) of $X_t$ (3.2) with $\theta_t$ varying. In the absence of ambiguity (i.e., $(\theta_t) = 0$), the firm considers a unique SDE (3.1) as the dynamics of the state variable. According to this setup, the firm's Knightian uncertainty is limited "only" to the drift of the driving process. It is a consequence of the Brownian environment and the assumption of equivalence between the measures in $\Pi$

---

[4]A completed natural filtration $\mathcal{F}_t$ is generated by $\sigma(W_s; s \leq t)$ and the $Q$-null sets of $\mathcal{F}_T$.

(Epstein and Wang, 1994).

## 3.2.2 Irreversible investment

The irreversible investment decision is an example where an agent decides whether to exercise an option (i.e., invest) in order to pursue upside potential. It is analogous to the exercising of an American call option written on the underlying project value with a fixed investment cost as the strike price. The model was first developed by McDonald and Siegel (1986). The following example is a continuous time extension from Miao and Wang (2011) and is a special case of the model in Nishimura and Ozaki (2007).

Denote the value of the option to invest by $V(x)$ where $x$ represents the underlying project value. In the absence of ambiguity, the process $(X_t)$ evolves according to (3.1). Suppose the firm must pay a fixed sunk cost $I$ to invest. Then, the payoff from investing at time $t$ is $\Omega(X_t) = X_t - I$ and the firm choosing the stopping policy $\tau := \inf\{t : X_t(\omega) = x^*\}$ maximizes its expected present value, given the current $X_t = x$,

$$V(x) = \max_{\tau} \ \mathbb{E}_t[e^{-\rho\,(\tau-t)}\Omega(X_\tau)]. \tag{3.3}$$

Note that $\rho$ is the firm's subjective discount rate. I allow the stopping time $\tau$ to be infinite, so that (3.3) may be time-homogenous.

During the small time interval $[t, t + dt)$, the firm's optimal investment decision may be described as choosing between two options: invest now at time $t$ or wait for a short time period $dt$. If the firm chooses to exercise the option at time $t$, it obtains the payoff $\Omega(x)$. If the firm waits for $dt$, the underlying project value becomes $X_t + dX_t$ where $dX_t$ is a random increment from the perspective of time $t$. The firm captures a capital gain $V(X_t + dX_t)$. Because the investment opportunity yields no cash flow up to the time $\tau$ that the investment is undertaken, the only gain from holding the option is its capital appreciation. To express this into time-$t$ equivalent units, it must be discounted by $e^{-\rho\,dt}$. Then, by the principle of optimality, the Bellman equation for the firm's optimal investment decision problem (3.3) becomes

$$V(x) = \max \left\{ \underbrace{\Omega(x)}_{\substack{\text{Payoff of} \\ \text{immediate investment}}} , \underbrace{e^{-\rho\,dt}\mathbb{E}_t[V(x + dx)]}_{\text{Capital gain}} \right\}$$

$$= \max \left\{ \Omega(x) , \mathbb{E}_t[dV] + V - \rho\,V\,dt \right\}. \tag{3.4}$$

In the last line, I use $e^{-\rho\,dt} \approx (1 - \rho dt)$ as $dt \to 0$ and Ito's lemma.

Now, suppose the firm faces Knightian uncertainty about the true dynamics of the underlying value process $(X_t)$. Assume the manager is ambiguity averse with max-min preferences (Gilboa and Schmeidler, 1989). Then, the firm's optimal investment decision is

$$
\begin{aligned}
V(x) &= \max\left\{\Omega(x)\,,\, \min_{Q^\theta \in \Pi} e^{-\rho\,dt}\mathbb{E}_t^{Q^\theta}[V(X_t + dX_t)]\right\} \\
&= \max\left\{\Omega(x)\,,\, \min_{Q^\theta \in \Pi} \mathbb{E}^{Q^\theta}[dV] + V - \rho V dt\right\}.
\end{aligned}
\tag{3.5}
$$

In order to ensure dynamic consistency in the recursive relation, I assume $i.i.d$ Knightian uncertainty: $\theta_t(\omega) \in [-\kappa, \kappa]$ for any given $t$ (Epstein and Wang, 1994; Nishimura and Ozaki, 2007).[5] In the time-homogenous case, choosing the optimal stopping time is equivalent to choosing the optimal threshold. Therefore, there exits a threshold $x^*$ such that when $x < x^*$, it is optimal for the firm to wait

$$
\begin{aligned}
\rho V &= \min_{Q^\theta \in \Pi} \frac{1}{dt}\mathbb{E}^{Q^\theta}[dV] \\
&= \min_{\theta_t \in [-\kappa,\kappa]} \left((\mu - \theta_t\sigma)xV'(x) + \frac{1}{2}\sigma^2 x^2 V''(x)\right).
\end{aligned}
\tag{3.6}
$$

Assuming that $V(\cdot)$ is twice differentiable and conjecture that $V'(x) > 0$. Then I have the following Hamilton-Jacobi-Bellman equation

$$
\rho V = (\mu - \kappa\sigma)xV'(x) + \frac{1}{2}\sigma^2 x^2 V''(x).
\tag{3.7}
$$

---

[5]$I.i.d$ ambiguity is the special case of strong rectangularity. The details are available in Epstein and Wang (1994) and Nishimura and Ozaki (2007). Rectangularity of $\Pi$ (i.e., $(\theta_t)$) is necessary for the dynamic consistency of the value function. Consider the following value function:

$$
V(x_t) = \min_{Q^\theta \in \Pi} \mathbb{E}_t^{Q^\theta}\left[\int_t^T e^{-\rho\,(s-t)}f(X_s)ds\right].
$$

Because of the rectangularity of $\Pi$, the above value function satisfies the following recursive relation:

$$
V(x_t) = \min_{Q^\theta \in \Pi} \mathbb{E}_t^{Q^\theta}\left[\int_t^{t+h} e^{-\rho\,(s-t)}f(X_s)ds + e^{-\rho h}V(X_{t+h})\right].
$$

That is, when $\xi$ is $\mathcal{F}_T$-measurable and $s \le t \le T$, the rectangularity of $\Pi$ ensures

$$
\min_{Q \in \Pi} \mathbb{E}_t^Q[\xi] = \min_{Q \in \Pi} \mathbb{E}_t^Q\left[\mathbb{E}_s^Q[\xi]\right] = \min_{Q \in \Pi} \mathbb{E}_t^Q\left[\min_{Q' \in \Pi} \mathbb{E}_s^{Q'}[\xi]\right].
$$

Moreover, $V(x)$ must satisfy the conditions:

$$V(0) = 0, \tag{3.8}$$

$$V(x^*) = \Omega(x^*) = x^* - I, \qquad \text{value matching} \tag{3.9}$$

$$V'(x^*) = \Omega'(x^*) = 1. \qquad \text{smooth pasting} \tag{3.10}$$

The first condition ensures when the value of project is zero, the option written on it must have no value. The second condition ensures the continuity of value function $V(x)$. The smooth pasting condition is necessary because $x^*$ is endogenously chosen by the manager of the firm. The details about smooth pasting condition are available in Appendix 3.A.1 on page 99 (or see Dixit and Pindyck (1994, Chapter 4)).

The following propositions 3.1 is special cases of the results in Nishimura and Ozaki (2007).

**Proposition 3.1.** *(a) The value function $V(x)$ that satisfies (3.7) and the conditions (3.8), (3.9), and (3.10) is:*

$$V(x) = A_1 x^{\beta_1},$$

*where*

$$A_1 = \frac{1}{\beta_1} \left( \frac{\beta_1}{\beta_1 - 1} I \right)^{1 - \beta_1}.$$

*In addition, the critical investment trigger $x^*$ is*

$$x^* = \frac{\beta_1}{\beta_1 - 1} I = \left( 1 + \frac{1}{\beta_1 - 1} \right) I > I, \tag{3.11}$$

*where $\beta_1 > 1$ is one of the solutions of the following characteristic equation associated with the HJB equation (3.7): $Q(\beta) = \frac{1}{2}\sigma^2 \beta(\beta - 1) + (\mu - \kappa\sigma)\beta - \rho = 0$. Specifically,*

$$\beta_1 = \frac{1}{\sigma^2} \left( -\left( \mu - \kappa\sigma - \frac{\sigma^2}{2} \right) + \sqrt{\left( \mu - \kappa\sigma - \frac{\sigma^2}{2} \right)^2 + 2\rho\sigma^2} \right).$$

*(b) In the absence of Knightian uncertainty (i.e., $\kappa = 0$), an increase in risk ($\sigma^2$) raises the critical value (exercise boundary) $x^*$ at which it is optimal for the firm to invest. That is,*

$$\frac{\partial x^*}{\partial \sigma^2} > 0.$$

*(c) An increase in ambiguity ($\kappa$) lowers the critical value $x^*$ at which it is optimal for the*
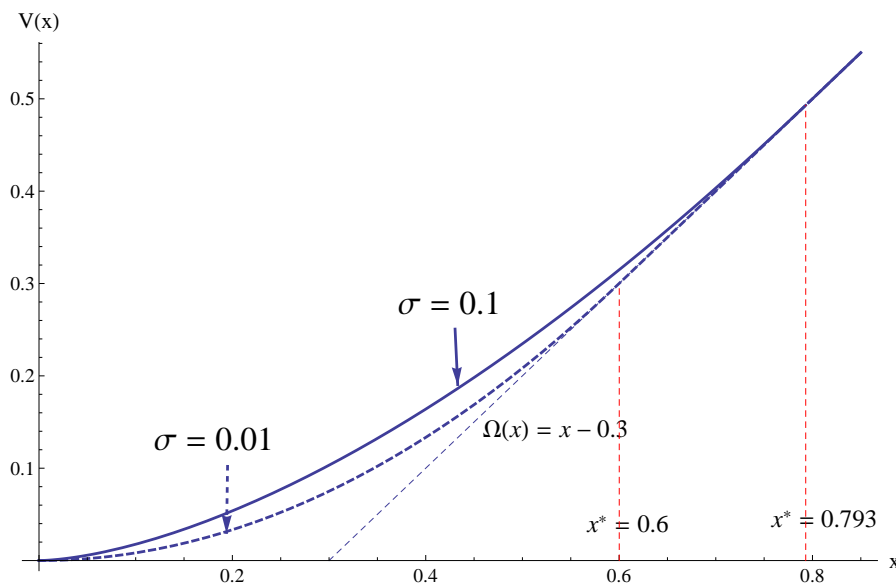
*firm to invest. That is,*

$$\frac{\partial x^*}{\partial \kappa} < 0.$$

*Moreover, when the firm confronts a sufficiently large amount of Knightian uncertainty, it is optimal for the firm to follow the NPV rule. That is,*

$$\lim_{\kappa \to \infty} x^* = I.$$
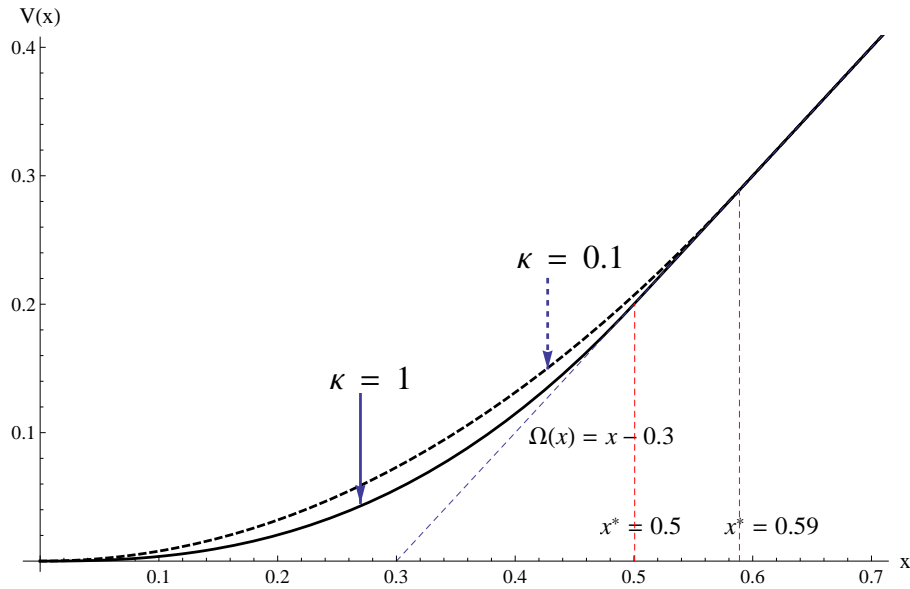
*Proof.* See Appendix 3.A.2 on page 99. □



This figure illustrates the optimal investment trigger $(x^*)$ when changing the level of risk of the underlying project value process. To generate this figure, I set the expected growth rate of project value as $\mu = 0.05$, the firm's subjective discount rate $\rho = 0.1 > \mu$, and the fixed sunk cost at the time of investment $I = 0.3$. Moreover, I assume that there is no uncertainty $(\kappa = 0)$.

**Figure 3.1 Comparative statistics of the optimal investment trigger $(x^*)$ with respect to the change in risk $(\sigma)$**

Figure 3.1 and Figure 3.2 provide numerical examples to illustrate Proposition 3.1-(b). As can be seen from Figure 3.1, when the level of risk $(\sigma)$ increases from 0.01 to 0.1, the optimal investment trigger $x^*$ increases from 0.6 to 0.793. That is, the firm is more likely to wait to capture the upside potential. Figure 3.2 illustrates that in contrast, the effect of uncertainty deceases the value of waiting. When the level of uncertainty $(\kappa)$ increases from 0.1 to 1, the optimal investment trigger $x^*$ decreases from 0.59 to 0.5. Therefore, when the firm faces the higher ambiguity about the expected growth rate of the underlying project value, it is less likely to wait. It is important to understand that Proposition 3.1-(b) crucially

relies on the assumption that there is no uncertainty after the investment. That is, $\Omega(x)$ involves no Knightian uncertainty.



This figure illustrates the optimal investment trigger $(x^*)$ when changing the level of Knightian uncertainty of the underlying project value process. To generate this figure, I set the expected growth rate of project value as $\mu = 0.05$, the level of risk of the project value process $\sigma = 0.01$, the firm's subjective discount rate $\rho = 0.1 > \mu$, and the fixed sunk cost at the time of investment $I = 0.3$.

**Figure 3.2 Comparative statistics of the optimal investment trigger $(x^*)$ with respect to the change in the level of Knightian uncertainty $(\kappa)$**

Standard option theory indicates that the simple NPV rule is incorrect $(x^* > I)$; risk and irreversibility drive a wedge between the critical value $x^*$ and the sunk cost $I$ (see Dixit and Pindyck (1994)). However, when the firm confronts a sufficiently large amount of Knightian uncertainty, it is optimal to follow the NPV rule. That is, as the amount of uncertainty $(\kappa)$ is infinite, the optimal investment trigger $x^*$ converges to $I$.

### 3.2.3   Exit option

This example is a continuous time extension of an example in Miao and Wang (2011). Consider a firm in an industry. The process $(X_t)$ can be interpreted as a productivity (or demand) shock. Staying in business at time $t$ yields profits with a fixed cost $c > 0$, so that the net profit is $f(x) = x - c$. The firm may decide to exit and seek outside option $\gamma$ that is assumed to be constant. Then the firm chooses the stopping policy $\tau = \inf\{t : X_t(\omega) = x^*\}$

to maximize, given $X_t = x$,

$$V(x) = \max_{\tau} \ \mathbb{E}_t\left[\int_t^{\tau} e^{-\rho(s-t)} f(X_s)ds + e^{-\rho(\tau-t)}\Omega(X_\tau)\right].\qquad(3.12)$$

Again I assume that the stopping time $\tau$ can be infinite, so that the value function $V(x)$ is time-homogenous.

Similar to the irreversible investment example, for the short time interval $[t, t + dt)$, the firm chooses either to exit and obtain $\Omega(x)$ or continue to obtain profits and the capital gain. Precisely, solving (3.12) is equivalent to solving the following Bellman equation, given $X_t = x$:

$$V(x) \ = \ \max\left\{\underbrace{\Omega(x)}_{\text{outside option}} \ , \ \underbrace{f(x)dt}_{\text{dividend}} + \underbrace{\mathbb{E}_t\left[e^{-\rho\,dt}V(X_{t+dt})\right]}_{\text{capital gain}}\right\}$$

$$= \ \max\left\{\Omega(x), f(x)dt + \mathbb{E}_t[dV] + V - \rho V dt\right\}.\qquad(3.13)$$

When the firm faces Knightian uncertainty about the dynamics of $(X_t)$, (3.13) becomes:

$$V(x) \ = \ \max\left\{\Omega(x), f(x)dt + \min_{Q^\theta \in \Pi} \mathbb{E}_t^{Q^\theta}[dV] + V - \rho V dt\right\}$$

When $x > x^*$, $V(x)$ satisfies

$$\rho V(x) \ = \ f(x) + \min_{Q^\theta \in \Pi} \frac{1}{dt}\mathbb{E}^{Q^\theta}[dV]$$

$$= \ f(x) + \min_{\theta_t \in [-\kappa,\kappa]} \left((\mu - \theta_t\sigma)xV'(x) + \frac{1}{2}\sigma^2 x^2 V''(x)\right).$$

where in the second line I use the *i.i.d.* uncertainty assumption. Conjecturing that $V'(x) > 0$, I have the following HJB equation:

$$\rho V = (\mu - \kappa\sigma)xV' + \frac{1}{2}\sigma^2 x^2 V'' + f(x),\qquad(3.14)$$

where $f(x) = x - c$. The solution of (3.14) is a linear combination of a general solution and a particular solution:

$$V(x) = v_p(x) + A_1 x^{\beta_1} + A_2 x^{\beta_2},$$

where $\beta_1 > 1$ and $\beta_2 < 0$ are the solutions of the characteristic equation $Q(x)$:

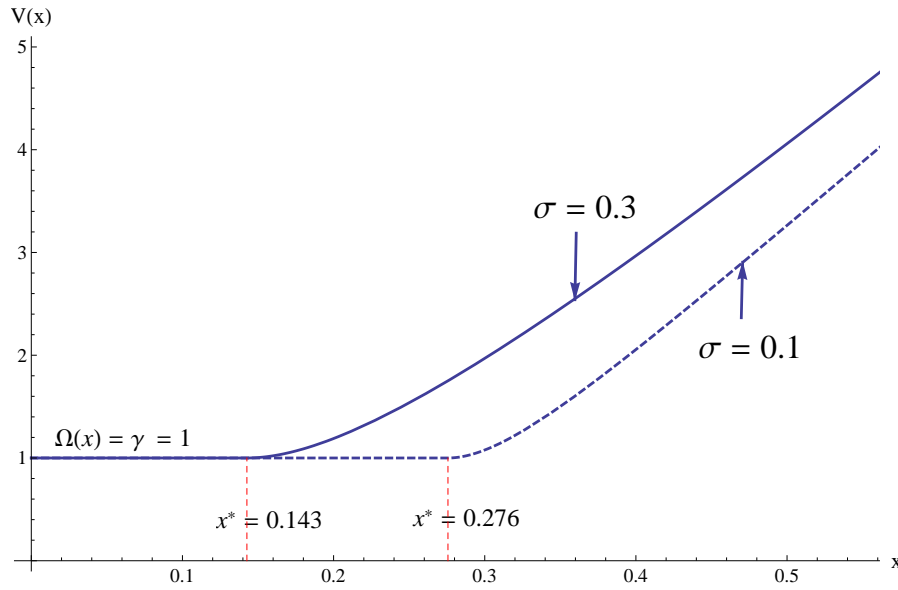$$Q(\beta) = \frac{1}{2}\sigma^2\beta(\beta - 1) + (\mu - \kappa\sigma)\beta - \rho = 0. \tag{3.15}$$

In addition, $V(x)$ must satisfy the following conditions:

$$\lim_{x\to\infty} V(x) = v_p(x) \qquad \text{transversality} \tag{3.16}$$

$$V(x^*) = \Omega(x^*) = \gamma \qquad \text{value matching} \tag{3.17}$$

$$V'(x^*) = \Omega'(x^*) = 0 \qquad \text{smooth pasting} \tag{3.18}$$

Note that the transversality condition ensures that $A_1 = 0$ because $\beta_1 > 1$.



This figure illustrates the optimal exit trigger $(x^*)$ when changing the level of risk of the underlying project value process. To generate this figure, I set the expected growth rate of project value as $\mu = 0.02$, the firm's subjective discount rate $\rho = 0.1 > \mu$, and the constant outside option value $\Omega(x) = \gamma = 1$. Moreover, I assume that there is no uncertainty $(\kappa = 0)$.

**Figure 3.3 Comparative statistics of the optimal exit trigger $(x^*)$ with respect to the change in risk $(\sigma)$**

**Proposition 3.2.** *(a) The value function $V(x)$ that satisfies* (3.14) *and the conditions* (3.16), (3.17), *and* (3.18) *is:*

$$V(x) = v_p(x) + A_2 x^{\beta_2},$$

*where*

$$v_p(x) = \frac{x}{\rho - (\mu - \kappa\sigma)} - \frac{c}{\rho}, \qquad A_2 = -\frac{1}{\beta_2}(x^*)^{1-\beta_2}\frac{1}{\rho - (\mu - \kappa\sigma)},$$

and the optimal exit trigger $x^*$ is

$$x^* = \frac{\beta_2}{\beta_2 - 1} \left( \gamma + \frac{c}{\rho} \right) (\rho - (\mu - \kappa\sigma)),$$

where $\beta_2 < 0$ are the solution of (3.15). Specifically,

$$\beta_2 = \frac{1}{\sigma^2} \left( -\left( \mu - \kappa\sigma - \frac{\sigma^2}{2} \right) - \sqrt{\left( \mu - \kappa\sigma - \frac{\sigma^2}{2} \right)^2 + 2\rho\sigma^2} \right).$$

(b) In the absence of Knightian uncertainty, an increase in risk lowers the optimal exit trigger. That is,

$$\frac{\partial x^*}{\partial \sigma^2} < 0.$$

(c) An increase in Knightian uncertainty ($\kappa$) raises the optimal exit trigger. In other words,

$$\frac{\partial x^*}{\partial \kappa} > 0.$$

*Proof.* See Appendix 3.A.3 on page 102. □

Figure 3.3 and Figure 3.4 provide numerical examples to illustrate Proposition 3.2-(b). As can be seen from Figure 3.3, when the firm confronts a higher level of risk, the continuation value increases while the outside option stays constant. Hence, the optimal exit trigger decreases. By contrast, Figure 3.4 illustrates that as the firm confronts a higher level of uncertainty, the continuation value decreases while the outside option stays const. Therefore, the optimal exit trigger increases with Knightian uncertainty.

According to Proposition 3.2, firms may stay in business longer in riskier situations, even though they suffer losses, hoping for the upside potential. However, as stated by Miao and Wang (2011), this prediction seems to be inconsistent with the empirical observation of the large amount of quick exit in the IT industry in recent years. Knightian uncertianty may provide a plausible explanation to this issue. When firms are more uncertain about the industry demand or their productivity, they are less sure about the likelihood of their effective cash flows. When the firm's uncertainty increases, the continuation value of firms is lower and the exit trigger higher. Therefore, it is optimal for firms to exit earlier.
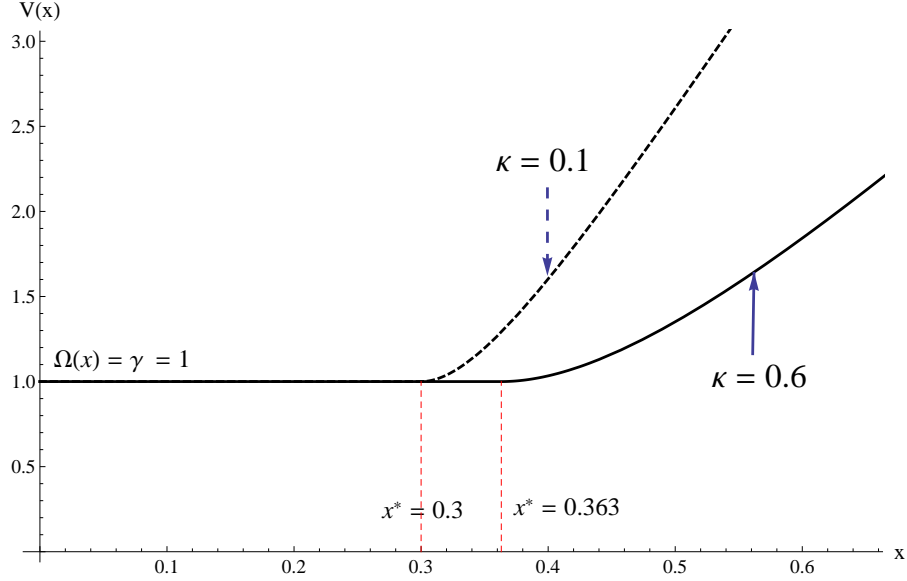
This figure illustrates the optimal exit trigger $(x^*)$ when changing the level of Knightian uncertainty of the underlying project value process. To generate this figure, I set the expected growth rate of project value as $\mu = 0.02$, the level of risk $\sigma = 0.1$, the firm's subjective discount rate $\rho = 0.1 > \mu$, the time-invariant production cost $c = 0.5$, and the constant outside option value $\Omega(x) = \gamma = 1$.

**Figure 3.4 Comparative statistics of the optimal exit trigger $(x^*)$ with respect to the change in the level of Knightian uncertainty $(\kappa)$**

## 3.3 Dynamic Capital Structure with Uncertainty

### 3.3.1 The model with the investors' uncertainty aversion

The standard model of dynamic capital structure model is based on Leland (1994). A firm's asset value is assumed to follow the risk-neutral dynamics

$$dV_t/V_t = \mu\, dt + \sigma\, dW_t \tag{3.19}$$

where $W_t$ is a Brownian motion under the risk-neutral measure $Q$, and $\mu < r$ is the expected asset growth rate. The firm generates cash flow at the rate of $\delta V_t$ at time $t$, for some constant $\delta \in (0, \infty)$.

The firm issues debt so as to obtain the tax shields offered for interest expense. It has a constant marginal tax rate $\tau \in [0, 1)$. For a simple time homogenous setting, the debt is modeled as a consol bond that pays coupons indefinitely at a constant rate $c$. Thus, in the interval $[t, t + dt)$, the coupon payment is $c\, dt$.

The investors of the firm (the equity and bond holders) are uncertain about the dynamics

90

of the asset value process $(V_t)$. Instead, they maintain a set of perturbed versions of reference motion $(V_t)$ that they think possible: they think there are the perturbed versions of $W_t$ given by:

$$dW_t^\theta = \theta_t dt + dW_t.$$

As discussed in Section 3.2.1, the evolution of the firm's asset process under the investors' uncertainty about the reference measure $Q$ is

$$\frac{dV_t}{V_t} = \mu dt + \sigma(dW_t^\theta - \theta_t dt) = (\mu - \theta_t \sigma)dt + \sigma dW_t^\theta, \tag{3.20}$$

and note that an element of the set of measures $\Pi = \{Q^\theta\}$ can be indexed by $(\theta_t)$. I assume that there is no information asymmetry between the claim holders: they face the same amount of uncertainty (i.e., share the same set of priors $\Pi$) and have the same max-min ambiguity aversion preferences.

The firm is operated by its shareholders. For simplicity, their only action is choosing when to default. A liquidation policy is a choice of optimal stopping time $\tau_d(v_B) = \inf\{t : V_t(\omega) = v_B\}$ where $v_B$ is an endogenous optimal default trigger. Following Duffie and Lando (2003), given an asset level at liquidation of $v_B$, I assume that the liquidation value of assets is the expected present value at the time $t$ of the cash flows to be generated by the assets:

$$\min_{Q^\theta \in \Pi} \mathbb{E}_t^{Q^\theta}\left[\int_t^\infty e^{-r(s-t)}\delta V_s ds\right] = \frac{\delta V_t}{(r - \mu + \kappa\sigma)}\bigg|_{V_t = v_B}. \tag{3.21}$$

At the chosen default time, a fraction $\alpha \in [0,1]$ of the assets are lost as a deadweight cost. The value of remaining assets

$$\frac{(1-\alpha)\delta v_B}{(r - \mu + \kappa\sigma)}$$

is assigned to the debt-holders by absolute priority.

The equity holders of the firm chooses the optimal stopping policy $\tau_d(v_B)$ to maximize, given the current value $V_t = v$,

$$E(v) = \max_{\tau_d} \min_{Q^\theta \in \Pi} \mathbb{E}_t^{Q^\theta}\left[\int_t^{\tau_d} e^{-r(s-t)}(\delta V_s - (1-\tau)c)ds\right]. \tag{3.22}$$

When the firm is solvent, the value of equity satisfies the following no-arbitrage equation

(i.e., HJB equation):

$$
\begin{aligned}
rE(v) &= \delta v - (1-\tau)c + \min_{Q^\theta \in \Pi} \frac{1}{dt} \mathbb{E}_t^{Q^\theta}[dE(v)] \\
&= \delta v - (1-\tau)c + \mu v E'(v) + \frac{1}{2}\sigma^2 v^2 E''(v) + \underbrace{\min_{Q^\theta \in \Pi} \frac{1}{dt} \mathbb{E}_t^{Q^\theta}[dW]}_{\substack{\neq 0 \\ \text{due to ambiguity}}} \sigma v E'(v) \\
&= \delta v - (1-\tau)c + \mu v E'(v) + \frac{\sigma^2}{2} v^2 E''(v) + \min_{Q^\theta \in \Pi} \frac{1}{dt}\mathbb{E}_t^{Q^\theta}[-\theta_t\,dt + dW^\theta]\sigma E'(v) \\
&= \delta v - (1-\tau)c + \mu v E'(v) + \frac{\sigma^2}{2} v^2 E''(v) + \min_{\theta_t \in [-\kappa,\kappa]}(-\theta_t)\,\sigma v E'(v) \\
&= \delta v - (1-\tau)c + \underbrace{(\mu - \kappa\sigma)}_{\substack{\text{Ambigiuty} \\ \text{adjusted drift}}} v E'(v) + \frac{\sigma^2}{2} v^2 E''(v) \tag{3.23}
\end{aligned}
$$

Here, I assume that from the set of possible pricing measures, the market chooses the minimum pricing measure in the set. In the third line, I assume $i.i.d.$ uncertainty and conjecture that $E'(v) > 0$. Limited liability for the equity holders ensures the following conditions:

$$
\lim_{v \to \infty} E(v) = \frac{\delta v}{r - \mu + \kappa\sigma}, \qquad E(v_B) = 0, \qquad E'(v_B) = 0 \tag{3.24}
$$

**Proposition 3.3.** *The value of equity that satisfies the no-arbitrage equation* (3.23) *with the conditions* (3.24) *is*

$$
E(v; v_B, c) = \frac{\delta v}{(r - \mu + \kappa\sigma)} - \frac{(1-\tau)c}{r}\left(1 - \left(\frac{v}{v_B(c)}\right)^{-\beta}\right) - \frac{\delta v_B(c)}{(r - \mu + \kappa\sigma)}\left(\frac{v}{v_B(c)}\right)^{-\beta}. \tag{3.25}
$$

*The optimal default trigger $v_B$ is*

$$
v_B(c) = \frac{(1-\tau)c}{\delta r}\frac{\beta}{(\beta+1)}(r - \mu + \kappa\sigma), \tag{3.26}
$$

*where*

$$
\beta(\kappa, \sigma) = \frac{1}{\sigma^2}\left(\left(\mu - \kappa\sigma - \frac{\sigma^2}{2}\right) + \sqrt{\left(\mu - \kappa\sigma - \frac{\sigma^2}{2}\right)^2 + 2r\sigma^2}\right) > 0. \tag{3.27}
$$

*Proof.* See Appendix 3.B.1 on page 105. □

The value of equity (3.25) is merely the market value of assets at time $t$ less the market value of after-tax (effective) coupon payment to debt holders before default, and less the

market value of the debt holders' claim to $v_B$ at the time of default.

**Corollary 3.1.** *Fix a coupon rate c. Then, (a) in the absence of Knightian uncertainty, the optimal boundary $v_B$ decreases with the level of risk of cash flow. That is,*

$$\frac{\partial v_B}{\partial \sigma} < 0.$$

*(b) $v_B$ increases with the level of uncertainty. That is,*

$$\frac{\partial v_B}{\partial \kappa} > 0.$$

*Proof.* This is an immediate corollary to Proposition 3.2 on page 88 by setting $\gamma = 0$ and $f(v) = \delta v - (1-\tau)c.$ □

The market value of debt is the present value of a claim that pays the constant coupon rate $c\,dt$ for $[t, t+dt)$ until $\tau_d$ and a liquidation payment $\alpha\delta v_B/(r-\mu+\kappa\sigma)$ when the firm goes bankrupt at $\tau_d$. Hence, the no-arbitrage equation for the value of debt, given the current asset value $V_t = v$, is

$$rD(v) = c + \min_{Q^\theta} \frac{1}{dt}\mathbb{E}_t^{Q^\theta}[dD(v)] = c + (\mu - \kappa\sigma)vD'(v) + \frac{1}{2}\sigma^2 v^2 D''(v). \tag{3.28}$$

Again, I assume *i.i.d* uncertainty and conjecture that $D'(v) > 0$. It must satisfy the conditions:

$$D(v_B) = \frac{(1-\alpha)\delta v_B}{(r-\mu+\kappa\sigma)}, \qquad \lim_{v\to\infty} D(v) = \frac{c}{r}. \tag{3.29}$$

**Proposition 3.4.** *The value of debt that satisfies the no-arbitrage equation* (3.28) *with the conditions* (3.29) *is*

$$D(v; v_B, c) = \frac{c}{r} \cdot \left(1 - \left(\frac{v}{v_B(c)}\right)^{-\beta}\right) + \frac{(1-\alpha)\,\delta\,v_B(c)}{(r-\mu+\kappa\sigma)} \cdot \left(\frac{v}{v_B(c)}\right)^{-\beta}, \tag{3.30}$$
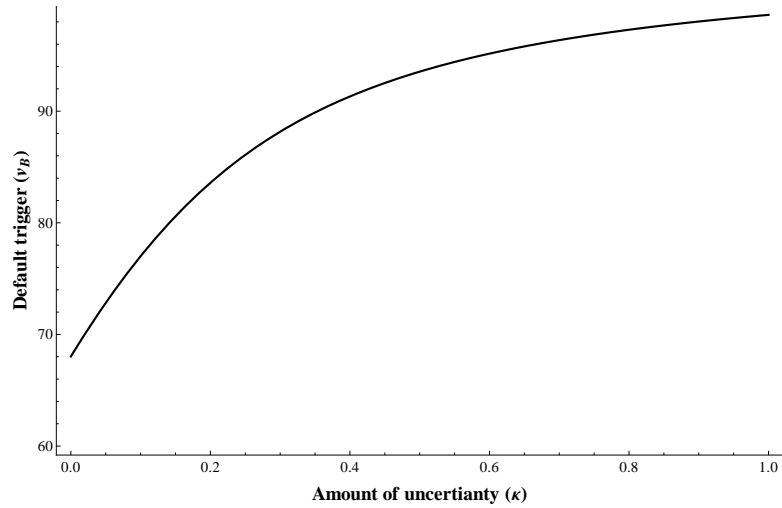
*where $\beta$ is given by* (3.27).

*Proof.* See Appendix 3.B.2 on page 105. □

The optimal coupon rate $c^*$, given the initial asset value $v$, is chosen so as to maximize the initial total firm value that is the initial value of equity (3.25) plus the value of debt (3.30).
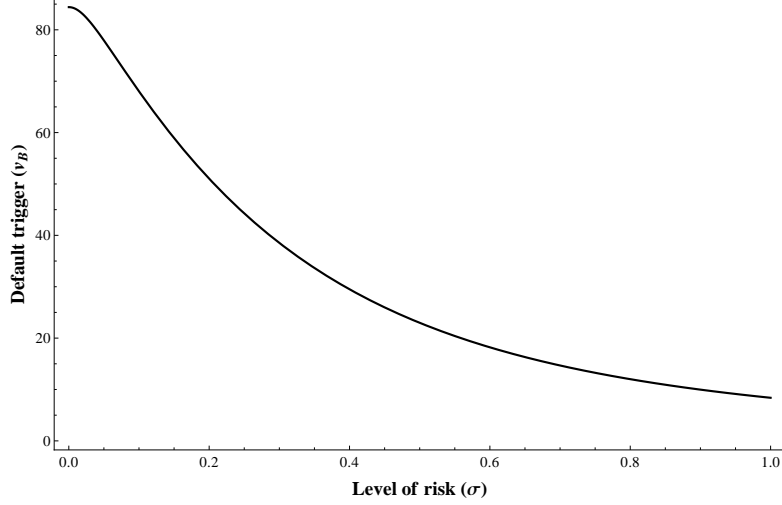
(A) This figure illustrates the comparative statics of the optimal trigger $v_B$ with respect to risk $\sigma$.
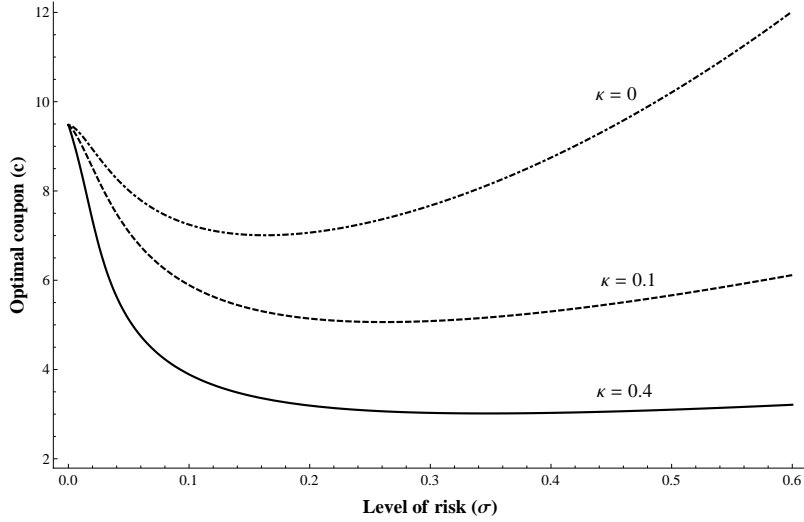


(B) This figure illustrates the comparative statics of the optimal trigger $v_B$ with respect to changes in uncertainty $\kappa$. To generate these figures, I fix the coupon rate $c = 8$, the expected growth asset rate $\mu = 0.0113$, the risk free discount rate $r = 0.06$, the marginal tax rate $\tau = 0.35$, the proportional deadweight cost $\alpha = 0.3$, the rate of cash flow $\delta = 0.05$ and the initial value of assets $v = 100$. Moreover, in Figure (A), I set $\kappa = 0$, while I set $\sigma = 0.1$ in Figure (B).

**Figure 3.5 Comparative statics of the optimal trigger $v_B$**
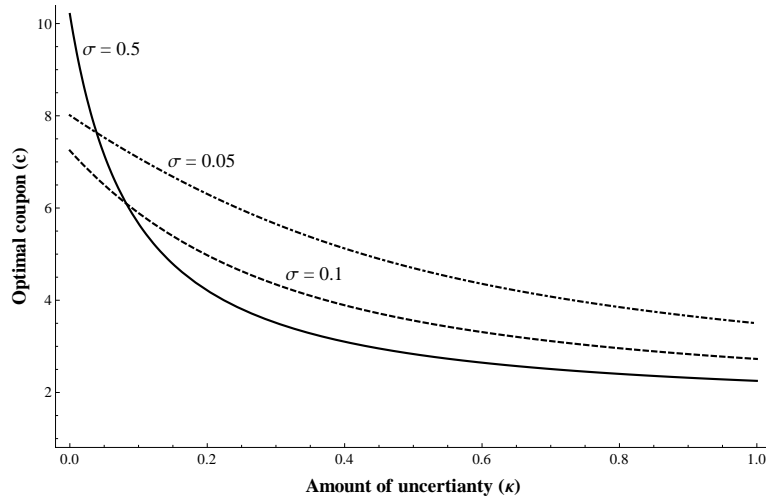
**Proposition 3.5.** *The optimal coupon rate is*

$$c^* = \frac{\delta v}{(r - \mu + \kappa\sigma)} \frac{r(1 + \beta)}{\beta(1 - \tau)} \left( \frac{(1 + \beta)\tau + \alpha(1 - \tau)\beta}{\tau} \right)^{(-1/\beta)} \tag{3.31}$$

*where $\beta$ is given by (3.27).*

(A) This figure illustrates the comparative statics of the optimal coupon $c$ with respect to risk $\sigma$. The dotted-dashed line represents when $\kappa = 0$ (without uncertainty), the dashed line when $\kappa = 0.1$, and the solid line when $\kappa = 0.6$.
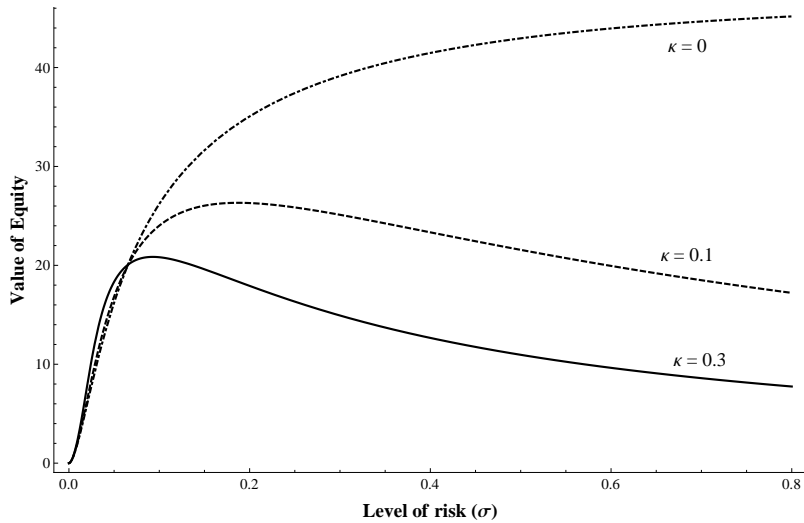


(B) This figure illustrates the comparative statics of the optimal coupon $c$ with respect to changes in uncertainty $\kappa$. The dotted-dashed line represents when $\sigma = 0.05$, the dashed line when $\sigma = 0.1$, and the solid line when $\sigma = 0.5$. To generate these figures, the other parameters are set as: the expected asset growth rate $\mu = 0.0113$, the risk free discount rate $r = 0.06$, the marginal tax rate $\tau = 0.35$, the proportional deadweight cost $\alpha = 0.3$, the rate of cash flow $\delta = 0.05$ and the initial value of assets $v = 100$.
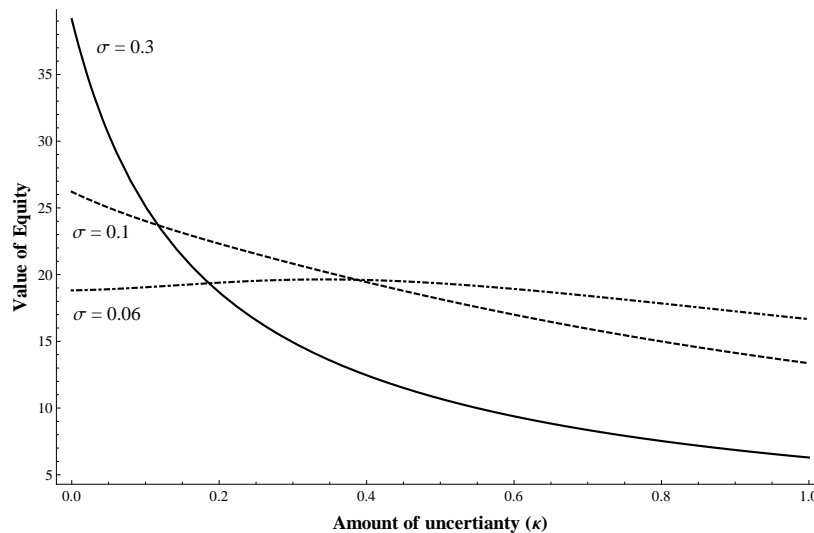
**Figure 3.6 Comparative statics of the optimal coupon $c^*$**

*Proof.* See Appendix 3.B.3 on page 106. □

*Remark* 3.1. For robustness, I verify that the sum of market value of equity given by (3.25) and equity given by (3.30) is the same as the unlevered market value of the firm plus the value of tax benefits and less the value of bankruptcy costs. I check this equivalence in

(A) This figure illustrates the comparative statics of the value of equity $E(v)$ with respect to changes in the level of risk ($\sigma$). The dotted-dashed line represents when $\kappa = 0$ (without uncertainty), the dashed line when $\kappa = 0.1$, and the solid line when $\kappa = 0.3$.



(B) This figure illustrates the comparative statics of the value of equity $E(v)$ with respect to changes in uncertainty $\kappa$. The dotted-dashed line represents when $\sigma = 0.06$, the dashed line when $\sigma = 0.1$, and the solid line when $\sigma = 0.3$. To generate these figures, the other parameters are set as: the expected asset growth rate $\mu = 0.0113$, the risk free discount rate $r = 0.06$, the marginal tax rate $\tau = 0.35$, the proportional deadweight cost $\alpha = 0.3$, the rate of cash flow $\delta = 0.05$ and the initial value of assets $v = 100$.

**Figure 3.7 Comparative statics of the value of equity $E(v)$**

Appendix 3.B.4 on page 106.

I turn to establish the comparative statics of the optimal leverage with respect to changes in the level of risk and uncertainty. For the numerical computations, I set the expected growth rate of the asset value process $\mu$, the level of risk $\sigma$, the initial value of asset $v$, the

risk free rate $r$, the marginal corporate tax rate $\tau$, the cash-flow generation rate $\delta$, and the proportional deadweight bankruptcy cost $\alpha$ as:

$$v = 100, \quad \delta = 0.05, \quad \mu = 0.011, \quad \sigma = 0.05, \quad r = 0.06, \quad \tau = 0.35, \quad \alpha = 0.3 \qquad (3.32)$$

The optimal leverage ratio can be computed as:

$$\text{Leverage ratio} = \frac{D(v; v_B(c^*), c^*)}{D(v; v_B(c^*), c^*) + E(v; v_B(c^*), c^*)} \qquad (3.33)$$

where $c^*$ is the optimal coupon found as (3.31), $E(v; v_B(c^*), c^*)$ is the market value of equity at $c^*$ using (3.25), and $D(v; v_B(c^*), c^*)$ is the market value of debt at $c^*$ using (3.30). Figure 3.8-(B) illustrates the optimal leverage ratio of the firm with respect to changes in the amount of uncertainty ($\kappa$). As can be seen, the optimal leverage ratio decreases with the amount of uncertainty. The benchmark model without uncertainty ($\kappa = 0$) predicts the optimal leverage ratio about 83%, given the level of risk $\sigma = 0.1$. When the amount of uncertainty increases as $\kappa = 0.8$, the optimal leverage decreases to about 68%. That is, the presence of substantial amount of Knightian uncertainty reduces the optimal leverage ratio by approximately 22%.

Figure 3.8-(A) illustrates the comparative statics of the firm's optimal leverage ratio with respect to changes in the level of risk. Qualitatively, the effect of risk also decreases the optimal leverage ratio. However, in order for risk alone to generate the low leverage ratio comparable to the level from the model with uncertainty, the required level of risk may increase up to $\sigma = 0.3$. In order for the model without uncertainty to generate 68% optimal leverage ratio, the required level of risk of cash flow should increase from $\sigma = 0.1$ (benchmark) to $\sigma = 0.3$.

## 3.4    Conclusion

Standard real option theory predicts that an increase in the risk of the underlying project value increases the value of waiting and delays the exercise of the option. Unlike risk, uncertainty erodes the value of waiting, so the firm exercises the option sooner. This prediction crucially depends on the assumption that there is no uncertainty about the termination payoff from the exercise of the option.

The benchmark dynamic capital structure model (Leland, 1994) generates an optimal
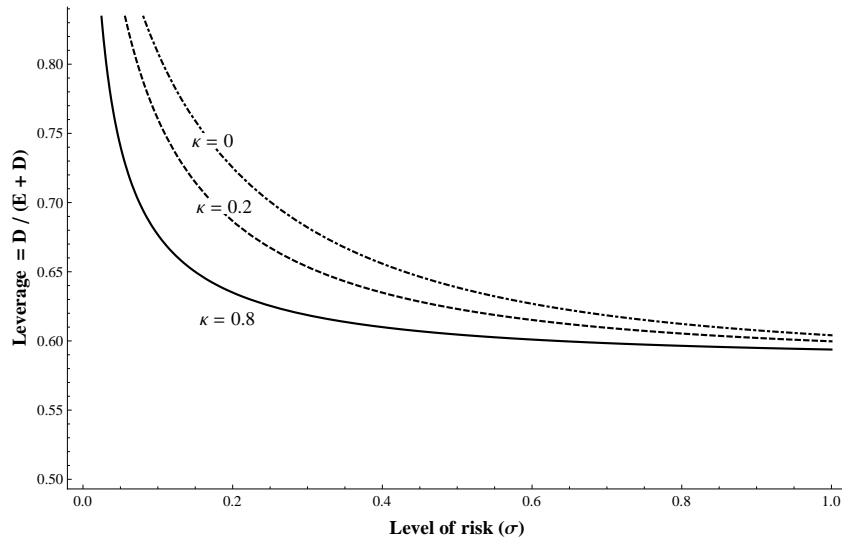
(A) This figure illustrates the optimal leverage ratio computed as (3.33) when changing the level of volatility of underlying process ($\sigma$). The dotted-dashed line represents when $\kappa = 0$ (without uncertainty), the dashed line when $\kappa = 0.2$, and the solid line when $\kappa = 0.8$.
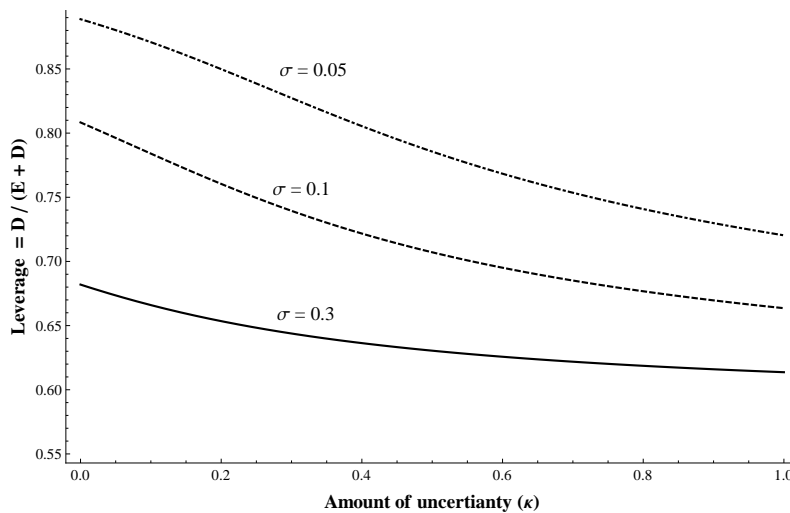


(B) This figure illustrates the optimal leverage ratio computed as (3.33) when changing the level of uncertainty ($\kappa$). The level of risk are fixed as $\sigma = 0.05$ (dot-dashed line), $\sigma = 0.1$ (dashed line), and $\sigma = 0.3$ (solid line). To generate these figures, the other parameters are set as: the expected asset growth rate $\mu = 0.0113$, the risk free discount rate $r = 0.06$, the marginal tax rate $\tau = 0.35$, the proportional deadweight cost $\alpha = 0.3$, the rate of cash flow $\delta = 0.05$ and the initial value of assets $v = 100$.

**Figure 3.8 Comparative statistics of the optimal leverage ratio**

leverage ratio for a firm of about 80%, given reasonable parameters. I incorporate investors' ambiguity about the expected growth rate of the firm's assets into the model. The ambiguity-augmented model generates an optimal leverage ratio of about 60%. This shows that uncertainty provides a plausible explanation for firms taking on low leverage even in a dynamic setting.

# 3.A  Proofs for Section 3.2

## 3.A.1  Proof of Smooth-Pasting Condition

*Proof.* For simplicity, let the continuation (waiting) region be $x < x^*$, and stopping region be $x > x^*$. By continuity of the value function, $V(x^*) = \Omega(x^*)$. Suppose the smooth pasting condition violated, then $V(x)$ and $\Omega(x)$ meet at a kink.

**case 1:** $V'(x^*) > \Omega'(x^*)$

It implies $V(x) < \Omega(x)$ when $x < x^*$. Therefore, it is optimal to stop and collect $\Omega(x)$ when $x < x^*$. However, it contradicts to the optimality of $x^*$.

**case 2:** $V'(x^*) < \Omega'(x^*)$

Instead of stopping at $x = x^*$, consider the alternative strategy that waits $\Delta t$ and reassess the situation. If the next state is $x^* + \Delta h$, stop, if the next state is $x^* - \Delta h$, continue. Let $p$ be $0.5(1 + \mu(x)/\sigma(x) \cdot \Delta h)$ where $\Delta h = \sqrt{\Delta t}$. By waiting $\Delta t$, the total payoff of the alternative strategy is,

$$
\begin{aligned}
& f(x^*)\Delta t + (1 - \rho\Delta t)\left[p\Omega(x^* + \Delta h) + (1-p)V(x^* - \Delta h)\right] \\
=\ & f(x^*)\Delta t + (1 - \rho\Delta t)\left[V(x^*) + (p\Omega'(x^*) - (1-p)V'(x^*))\Delta h + \cdots\right] \\
\stackrel{\Delta t \to 0}{=}\ & \Omega(x^*) + 0.5\left[\Omega'(x^*) - V'(x^*)\right]\Delta h > \Omega(x^*).
\end{aligned}
$$

In the second line, I use the value matching condition $V(x^*) = \Omega(x^*)$. Thus, the alternative strategy is better than stopping at $x = x^*$. Hence, it contradicts to the optimality of $x^*$.  $\square$

## 3.A.2  Proof of Proposition 3.1

*Proof.* It is well known that the solution of (3.7) is

$$
V(x) = A_1 x^{\beta_1} + A_2 x^{\beta_2}
$$

where are $\beta_1 > 0$ and $\beta_2 < 0$ are the solutions of the characteristic equation (3.15). Under the condition $\mu - \kappa\sigma < \rho$, it is immediate that $\beta_1 > 1$, To see this

$$
\begin{aligned}
\beta_1 &= \frac{1}{\sigma^2}\left(-\left(\mu - \kappa\sigma - \frac{\sigma^2}{2}\right) + \sqrt{\left(\mu - \kappa\sigma - \frac{\sigma^2}{2}\right)^2 + 2\rho\sigma^2}\right) \\
&> \frac{1}{\sigma^2}\left(-\left(\mu - \kappa\sigma - \frac{\sigma^2}{2}\right) + \left|\mu - \kappa\sigma + \frac{1}{2}\sigma^2\right|\right) \geq 1
\end{aligned}
$$

if $\mu - \kappa\sigma \geq -\frac{\sigma^2}{2}$, equal to unity. Otherwise, it is more than unity. And,

$$
\beta_2 = \frac{1}{\sigma^2}\left(-\left(\mu - \kappa\sigma - \frac{\sigma^2}{2}\right) - \sqrt{\left(\mu - \kappa\sigma - \frac{\sigma^2}{2}\right)^2 + 2\rho\sigma^2}\right) < 0. \tag{3.34}
$$

Combining $\beta_2 < 0$ and $V(0) = 0$ implies $A_2 = 0$. Hence,

$$
V(x_t) = A_1 x_t^{\beta_1}
$$

where $A_1$ and $x^*$ are determined by the value matching and smooth pasting conditions:

$$
\begin{aligned}
A_1 x^{*\beta_1} &= x^* - I, \\
A_1 \beta_1 x^{*\beta_1 - 1} &= 1.
\end{aligned}
$$

Specifically,

$$
x^* = \frac{\beta_1}{\beta_1 - 1}I = \left(1 + \frac{1}{\beta_1 - 1}\right)I > I, \qquad \text{such that} \quad \frac{\partial x^*}{\partial \beta_1} < 0
$$

$$
A_1 = \frac{1}{\beta_1}\left(\frac{\beta_1}{\beta_1 - 1}I\right)^{\beta_1 - 1}.
$$

With no uncertainty $(\kappa = 0)$, an increase of risk in the sense of mean preserving spread increase the optimal boundary. To see this,

$$
\partial Q = \frac{\partial Q}{\partial \beta}\frac{\partial \beta}{\partial \sigma^2} + \left.\frac{\partial Q}{\partial \sigma^2}\right|_{\beta = \beta_1 > 1} = 0,
$$

and we know

$$
\frac{\partial Q}{\partial \sigma^2} = \left.\beta(\beta - 1)\right|_{\beta = \beta_1} > 0 \qquad \left.\frac{\partial Q}{\partial \beta}\right|_{\beta = \beta_1} > 0.
$$

Hence, it has to be
$$\frac{\partial \beta_1}{\partial \sigma^2} < 0.$$

Therefore, the boundary increases with risk, because

$$\frac{\partial x^*}{\partial \sigma^2} = \frac{\partial x^*}{\partial \beta_1} \frac{\partial \beta_1}{\partial \sigma^2} > 0.$$

When $x < x^*$, the value of waiting increases with risk as well. To see this, when $x < x^*$,

$$
\begin{aligned}
\frac{\partial \log V(x)}{\partial \beta_1} &= \frac{\partial}{\partial \beta_1} \log\left((\beta_1 - 1)^{\beta_1 - 1} I^{1-\beta_1} \beta_1^{-\beta_1} x^{\beta_1}\right) \\
&= -\log I + \log(\beta_1 - 1) - \log \beta_1 + \log x \\
&< -\log I + \log(\beta_1 - 1) - \log \beta_1 + \log x^* \\
&= -\log I + \log(\beta_1 - 1) - \log \beta_1 + \log\left(\frac{\beta_1}{\beta_1 - 1} I\right) = 0.
\end{aligned}
$$

Because $V(x) > 0$ for all $x$, we have

$$\frac{\partial V(x)}{\partial \beta_1} < 0.$$

Therefore,

$$\frac{\partial V(x)}{\partial \sigma^2} = \frac{\partial V(x)}{\partial \beta_1} \frac{\partial \beta_1}{\partial \sigma^2} > 0.$$

On the other hand, fixing risk $(\sigma)$, an increase in uncertainty $(\kappa)$ lowers the optimal boundary. To see this,

$$\frac{\partial \beta_1}{\partial \kappa} = \frac{1}{\sigma}\left(1 - \frac{\left(\mu - \kappa\sigma - \frac{1}{2}\sigma^2\right)}{\sqrt{\left(\mu - \kappa\sigma - \frac{\sigma^2}{2}\right)^2 + 2\rho\sigma^2}}\right) > 0.$$

Hence,

$$\frac{\partial x^*}{\partial \kappa} = \frac{\partial x^*}{\partial \beta_1} \frac{\partial \beta_1}{\partial \kappa} < 0.$$

Also, the value of postponing investment is decreasing with uncertainty, because for $x < x^*$

$$\frac{\partial V(x)}{\partial \kappa} = \frac{\partial V}{\partial \beta_1} \frac{\partial \beta_1}{\partial \kappa} < 0.$$

$\square$

## 3.A.3  Proof of Proposition 3.2

*Proof.* The following particular solution satisfies (3.14):

$$v_p(x) = \frac{x}{\rho - (\mu - \kappa\sigma)} - \frac{c}{\rho}.$$

Now I turn to find the coefficients $A_1$ and $A_2$ of the general solution. As $x_t$ is sufficient large, the shutdown option has no value. That is,

$$\lim_{x \to \infty} V(x) = v_p(x).$$

Given the positivity of $\beta_1 > 1$, $A_1 = 0$. Hence,

$$V(x) = \frac{x}{\rho - (\mu - \kappa\sigma)} - \frac{c}{\rho} + A_2 x^{\beta_2},$$

where

$$\beta_2 = -\left(\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right) - \sqrt{\left(\frac{\delta}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2\rho}{\sigma^2}} < 0.$$

The coefficient $A_2$ and the optimal exit trigger $x^*$ can be determined by the value matching and smooth pasting conditions at the boundary $x^*$:

$$A_2 x^{*\beta_2} + \frac{x^*}{\rho - (\mu - \kappa\sigma)} - \frac{c}{\rho} \;=\; \gamma \qquad \text{by the value matching } V(x^*) = \Omega(x^*)$$

$$\beta_2 A_2 x^{*\beta_2 - 1} + \frac{1}{\rho - (\mu - \kappa\sigma)} \;=\; 0 \qquad \text{by the smooth pasting } V'(x^*) = \Omega'(x^*).$$

Then, we obtain

$$x^* = \frac{\beta_2}{\beta_2 - 1}\left(\gamma + \frac{c}{\rho}\right)(\rho - (\mu - \kappa\sigma)) \qquad \text{such that} \qquad \frac{\partial x^*}{\partial \beta_2} < 0,$$

$$A_2 = -\frac{1}{\beta_2} x^{*1 - \beta_2} \frac{1}{\rho - (\mu - \kappa\sigma)}.$$

When $\kappa = 0$, because of $\beta_2 < 0$, we know

$$\frac{\partial Q}{\partial \sigma^2} = \beta(\beta - 1)\Big|_{\beta = \beta_2} > 0, \qquad \frac{\partial Q}{\partial \beta}\Big|_{\beta = \beta_2} < 0.$$

Since

$$\partial Q = \frac{\partial Q}{\partial \beta} \frac{\partial \beta}{\partial \sigma^2} + \frac{\partial Q}{\partial \sigma^2}\bigg|_{\beta = \beta_2} = 0,$$

where $Q(\cdot)$ is the characteristic equation (3.15), we have the following

$$\frac{\partial \beta_2}{\partial \sigma^2} > 0.$$

Therefore,

$$\frac{\partial x^*}{\partial \sigma^2} = \frac{\partial x^*}{\partial \beta_2} \frac{\partial \beta_2}{\partial \sigma^2} < 0.$$

In the meantime, with no uncertainty $\kappa = 0$,

$$\frac{\partial V(x)}{\partial \beta_2} = \frac{\partial}{\partial \beta_2} \left( \frac{x}{\rho - \mu} + \frac{c}{\rho} + A_2 x^{\beta_2} \right).$$

Since the first two terms are not related with $\sigma$, it is enough to determine the sign of the last term. For $x \geq x^*$,

$$
\begin{aligned}
\frac{\partial \log A_2 x^{\beta_2}}{\partial \beta_2} &= \frac{\partial}{\partial \beta_2} \left( -\log(-\beta_2) + (1 - \beta_2) \log x^* - \log(\rho - \mu) + \beta_2 \log x \right) \\
&\geq \frac{\partial}{\partial \beta_2} \left( -\log(-\beta_2) + (1 - \beta_2) \log x^* - \log(\rho - \mu) + \beta_2 \log x^* \right) \\
&= \frac{\partial}{\partial \beta_2} \left( -\log(-\beta_2) + \log x^* - \log(\rho - \mu) \right) \\
&= \frac{\partial}{\partial \beta_2} \left( -\log(-\beta_2) + \log \left( \frac{\beta_2}{\beta_2 - 1} \left( \gamma + \frac{c}{\rho} \right) (\rho - \mu) \right) - \log(\rho - \mu) \right) \\
&= \frac{\partial}{\partial \beta_2} \left( -\log(1 - \beta_2) + \log(\gamma + c/\rho) \right) \\
&= \frac{1}{1 - \beta_2} > 0.
\end{aligned}
$$

Because $V(x) > 0$ for all $x$, therefore,

$$\frac{\partial V(x)}{\partial \sigma^2} = \frac{\partial V}{\partial \beta_2} \frac{\partial \beta_2}{\partial \sigma^2} > 0.$$

Now, I turn to investigate the effect of changes in uncertainty ($\kappa$) on the optimal investment trigger $x^*$, given the level of risk.

$$\frac{\partial x^*}{\partial \kappa} = \frac{\beta_2}{\beta_2 - 1} \left( \gamma + \frac{c}{\rho} \right) \sigma - \frac{1}{(\beta_2 - 1)^2} \left( \gamma + \frac{c}{\rho} \right) (\rho - \delta) \frac{\partial \beta_2}{\partial \kappa}, \tag{3.35}$$

where
$$0 < \frac{\partial \beta_2}{\partial \kappa} = \frac{1}{\sigma}\left(1 + \frac{\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}}{\sqrt{\frac{2\rho}{\sigma^2} + \left(\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right)^2}}\right) < \frac{2}{\sigma}.$$

Since
$$Q(\beta) = \beta^2 + \frac{2}{\sigma^2}(\mu - \kappa\sigma - \frac{1}{2}\sigma^2)\beta - \frac{2\rho}{\sigma^2} = 0,$$

and $\beta_2 < 0 < 1 < \beta_1$, we have

$$\beta_1 + \beta_2 = -2\left(\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right), \qquad \beta_1\beta_2 = -2\frac{\rho}{\sigma^2},$$

and
$$(\beta_1 - 1)(\beta_2 - 1) = \frac{2}{\sigma^2}(\mu - \kappa\sigma - \rho).$$

Plugging them into (3.35) on the previous page and with some computations, we have

$$\begin{aligned}
\frac{\partial x^*}{\partial \kappa} &= \left(\gamma + \frac{c}{\rho}\right)\left(\beta_2\sigma\frac{\sigma^3(\beta_1 - 1)}{2(\mu - \kappa\sigma - \rho)} - \frac{\sigma^4(\beta_1 - 1)^2}{4(\rho - \mu + \kappa\sigma)^2}(\rho - \mu + \kappa\sigma)\frac{\partial \beta_2}{\partial \kappa}\right) \\
&= \left(\gamma + \frac{c}{\rho}\right)\frac{\sigma^3(\beta_1 - 1)}{2(\rho - \mu + \kappa\sigma)}\left(-\beta_2 - \frac{\sigma}{2}(\beta_1 - 1)\frac{\partial \beta_2}{\partial \kappa}\right) \\
&= \left(\gamma + \frac{c}{\rho}\right)\frac{\sigma^3(\beta_1 - 1)}{2(\rho - \mu + \kappa\sigma)}\left(-\beta_2 - \frac{\sigma}{2}(\beta_1 - 1)\frac{1}{\sigma}\left(1 + \frac{\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}}{\sqrt{\frac{2\rho}{\sigma^2} + \left(\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right)^2}}\right)\right).
\end{aligned}$$

Let $A := \left(\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right)$ and $B := \frac{2\rho}{\sigma^2}$. Then,

$$\begin{aligned}
\frac{\partial x^*}{\partial \kappa} &= \underbrace{\left(\gamma + \frac{c}{\rho}\right)\frac{\sigma^3(\beta_1 - 1)}{2(\rho - \mu + \kappa\sigma)}}_{:=\Delta > 0}\left(A + \sqrt{A^2 + B} - \frac{1}{2}(-A + \sqrt{A^2 + B} - 1)\left(1 + \frac{A}{\sqrt{A^2 + B}}\right)\right) \\
&= \Delta\frac{1}{2}\left(1 + \frac{A}{\sqrt{A^2 + B}}\right)\left(A + \sqrt{A^2 + B} + 1\right) \\
&= \Delta\frac{1}{2}\left(1 + \frac{A}{\sqrt{A^2 + B}}\right)\left(\left(\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right) + \sqrt{\left(\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right)^2 + B + 1}\right) \\
&= \Delta\frac{1}{2}\left(1 + \frac{A}{\sqrt{A^2 + B}}\right)\left(\left(\frac{\mu - \kappa\sigma}{\sigma^2} + \frac{1}{2}\right) + \sqrt{\left(\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right)^2 + B}\right) \\
&> \Delta\frac{1}{2}\left(1 + \frac{A}{\sqrt{A^2 + B}}\right)\left\{\left(\frac{\mu - \kappa\sigma}{\sigma^2} + \frac{1}{2}\right) + \left|\frac{\mu - \kappa\sigma}{\sigma^2} - \frac{1}{2}\right|\right\} > 0.
\end{aligned}$$

$\square$

## 3.B  Proofs for Section 3.3

### 3.B.1  Proof of Proposition 3.3

*Proof.* The solution to (3.23) on page 92 is the sum of a particular solution and a general solution:

$$E(v) = \frac{\delta}{(r - \mu + \kappa\sigma)}v - \frac{(1 - \tau)}{r}c + A_2 v^{-\beta}, \tag{3.36}$$

where $-\beta$ is the positive solution of the characteristic equation associated with the HJB (3.23) on page 92:

$$Q(\beta) = \frac{1}{2}\sigma^2\beta(\beta - 1) + (\mu - \kappa\sigma)\beta - r = 0.$$

Solve it for $\beta$, then we have (3.27) on page 92. Apply the conditions (3.24) on page 92 and we have,

$$\frac{\delta}{r - \mu + \kappa\sigma}v_B - \frac{(1 - \tau)c}{r} + A_2 v_B^{-\beta} = 0$$

$$\frac{\delta}{r - \mu + \kappa\sigma} - A_2\beta v_B^{-\beta-1} = 0$$

Solve them for $v_B$, we have (3.26) on page 92 and

$$A_2 = \frac{(1 - \tau)c}{r}v_B - \frac{\delta}{(r - \mu + \kappa\sigma)}v_B^{\beta+1}$$

Plug it into (3.36), we have (3.25) on page 92. □

### 3.B.2  Proof of Proposition 3.4

*Proof.* The value of debt $D(v)$ that satisfies (3.28) on page 93 can be represented as

$$D(v) = \frac{c}{r} + A v^{-\beta}. \tag{3.37}$$

Apply the first boundary condition of (3.29) on page 93,

$$A = -\frac{c}{r}\left(\frac{v}{v_B}\right)^{-\beta} + \frac{(1 - \alpha)\delta\, v_B}{(r - \mu + \kappa\sigma)}\left(\frac{v}{v_B}\right)^{-\beta}.$$

Plug it into (3.37), we have (3.30) on page 93. □

### 3.B.3 Proof of Proposition 3.5

*Proof.* The initial share holders chooses the optimal coupon rate to maximize

$$\max_c E(v; v_B, c) + D(v; v_B, c)$$

where $E(v; v_B, c)$ is given by (3.25) and $D(v; v_B, c)$ is given by (3.30). From the FOC of the above, and after some manipulations, we have

$$v_B = v \left( \frac{\alpha(1 - \tau)\beta + \tau(1 + \beta)}{\tau} \right)^{-1/\beta}.$$

Plug it into (3.26) on page 92 and we have (3.31) on page 94. □

### 3.B.4 Balance sheet checks

The tax benefits of the firm are considered as the present value of a claim that pays $\tau c$ until the firm defaults, and pays nothing at the time default $\tau_d$.

**Proposition 3.6.** *Assume i.i.d. uncertainty and conjecture that $T'(v) > 0$. The tax benefits of the firm satisfies the following no-arbitrage equation*

$$r\, T(v) = \tau\, c + \min_{Q^\theta} \frac{1}{dt} \mathbb{E}_t^{Q^\theta}[dT(v)] = \tau\, c + (\mu - \kappa\sigma)T'(v) + \frac{1}{2}\sigma^2 v^2 T''(v),$$

*with the conditions:*
$$T(v_B) = 0, \qquad \lim_{v \to \infty} T(v) = \frac{\tau c}{r}. \tag{3.38}$$

*Therefore,*
$$T(v; v_B, c) = \frac{\tau c}{r}\left(1 - \left(\frac{v}{v_B}\right)^{-\beta}\right). \tag{3.39}$$

*Proof.* The tax benefits yields a flow of payoff $\tau c$. By the condition $TB(v_B) = 0$,

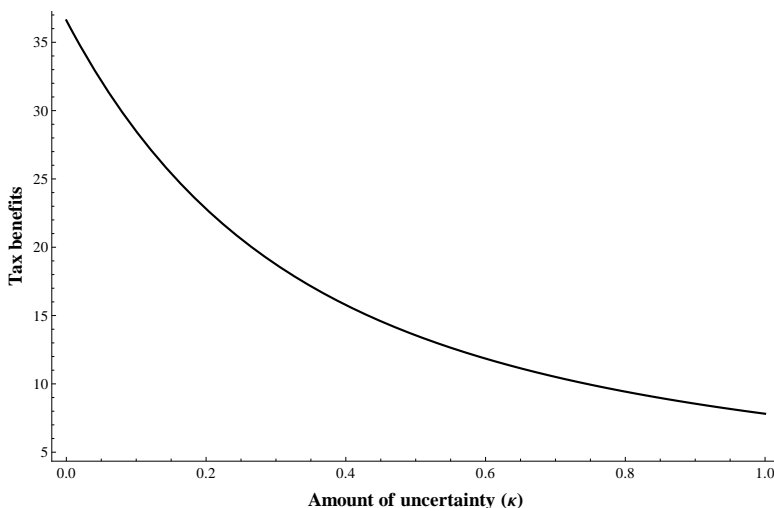$$TB(v_B) = \frac{\tau c}{r} + A_2 v_B^{-\beta} = 0$$

. Solve it for $A_2$,

$$A_2 = -\frac{\tau c}{r}\left(\frac{1}{v_B}\right)^{-\beta}.$$

Hence,

$$TB(v) = \frac{\tau c}{r} \left( 1 - \left( \frac{v}{v_B} \right)^{-\beta} \right).$$

□

Figure 3.9 provide the comparative statics of tax benefits with respect of changes in the amount of uncertainty $\kappa$. As can be seen, the tax benefits decrease with the amount of uncertainty.



This figure illustrates the tax benefits of the firm when changing the level of uncertainty $\kappa$. To generate this figure, I set the expected rate of return $\mu = 0.0113$, the risk free discount rate $r = 0.06$, the marginal tax rate $\tau = 0.35$, the proportional deadweight cost $\alpha = 0.3$, the rate of cash flow $\delta = 0.05$, the level of risk $\sigma = 0.1$, and the initial value of assets $v = 100$.

**Figure 3.9 Comparative statistics of the tax benefits with respect to changes in ambiguity ($\kappa$)**

The present value of bankruptcy cost is the present value of a claim that pays nothing until default, but when the firm defaults, it receives (negative payment) $\alpha \delta v_B / (r - \mu + \kappa \sigma)$ where $\alpha$ is the proportional bankruptcy cost.

The no-arbitrage equation for the bankruptcy costs is

$$
\begin{aligned}
rC(v) &= \max_{Q^\theta} \frac{1}{dt} \mathbb{E}_t^{Q^\theta}[dC(v)] \\
&= \max_{\theta_t \in [-\kappa, \kappa]} (\mu - \theta_t \sigma)\, v \underbrace{C'(v)}_{<0} + \frac{1}{2}\sigma^2 v^2 C''(v) \\
&= (\mu - \kappa\sigma)vC'(v) + \frac{1}{2}\sigma^2 v^2 C''(v).
\end{aligned}
\tag{3.40}
$$

107

In the first line, the worst case is the maximum appreciation because the deadweight cost will be lost at the time of default. In the second line, I assume the *i.i.d.* uncertainty and conjecture that the cost function is decreasing in $v$ (i.e., $C'(v) < 0$). It also must satisfy the conditions:

$$C(v_B) = \alpha \frac{\delta v_B}{(r - \mu + \kappa\sigma)}, \qquad \lim_{v \to \infty} C(v) = 0. \tag{3.41}$$

**Proposition 3.7.** *The market value of bankruptcy cost that satisfy no-arbitrage condition (3.40) and the conditions (3.41),*

$$C(v; v_B, c) = \alpha \frac{\delta v_B}{(r - \mu + \kappa\sigma)} \left(\frac{v}{v_B}\right)^{-\beta}. \tag{3.42}$$

*Proof.* By the transversality condition,

$$C(v) = A_2 v^{-\beta}$$

Use the condition $C(v_B) = \alpha \frac{\delta v_B}{(r-\mu+\kappa\sigma)}$,

$$A_2 = \frac{\alpha\,\delta v_B}{(r - \mu + \kappa\sigma)} \left(\frac{v}{v_B}\right)^{-\beta}$$

Plug it into the above, we have (3.42). $\qquad\square$

Figure 3.10 illustrates the comparative statics of tax benefits with respect of changes in the amount of uncertainty $\kappa$. As can be seen, the bankruptcy costs increase with the amount of uncertainty.
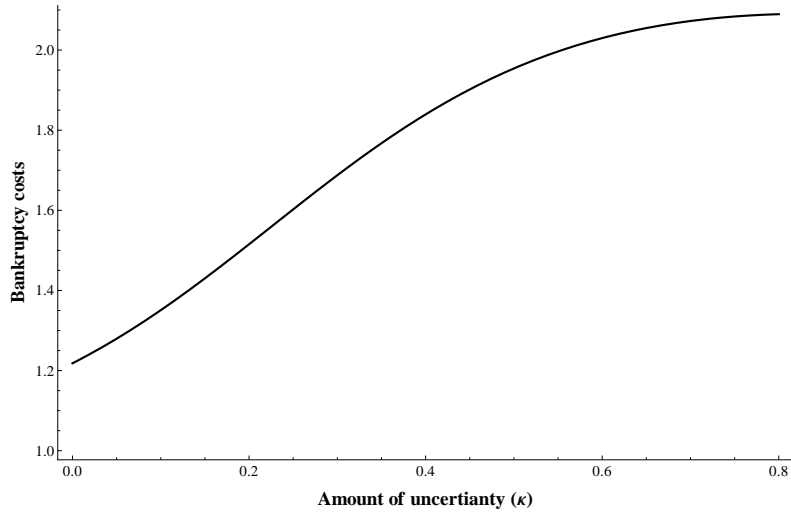
The total levered firm value is

$$\frac{\delta v}{(r - \mu + \kappa\sigma)} + T(v; v_B, c) - C(v; v_B, c), \tag{3.43}$$

that is the market value of cash flows to be generated by the current asset level $v$ plus the tax benefits of debt, and less the bankruptcy costs of debt. It must be the same as the right hand side of the balance sheet:

$$E(v; v_B, c) + B(v; v_B, c) \tag{3.44}$$

that is the sum of the market value of equity and debt. Hence, the market value of equity

This figure illustrates the bankruptcy cost of the firm when changing the level of uncertainty $\kappa$. To generate this figure, I set the expected rate of return $\mu = 0.0113$, the risk free discount rate $r = 0.06$, the marginal tax rate $\tau = 0.35$, the proportional deadweight cost $\alpha = 0.3$, the rate of cash flow $\delta = 0.05$, the level of risk $\sigma = 0.05$, and the initial value of assets $v = 100$.

**Figure 3.10 Comparative statistics of the bankruptcy cost with respect to the change in ambiguity ($\kappa$)**

can be found as

$$E(v; v_B, c) = \frac{\delta v}{(r - \mu + \kappa\sigma)} + T(v; v_B, c) - C(v; v_B, c) - B(v; v_B, c).$$

It is not surprising to see the above coincides to (3.25) on page 92. Note that the optimal coupon can be also found from the first order condition of (3.43),

$$\frac{\partial T(v; v_B, c)}{\partial c} = \frac{\partial C(v; v_B, c)}{\partial c}.$$

It is simple trade-off theory.

# Bibliography

Anderson, Keith B. (1993), 'Agency discretion or statutory direction: Decision making at the U.S. international trade commission', *Journal of Law and Economics* **36**(2), pp. 915–935.

Berk, Jonathan and Peter DeMarzo (2008), *Corporate Finance*, Prentice Hall.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2004), 'How much should we trust differences-in-differences estimates?', *The Quarterly Journal of Economics* **119**(1), 249–275.

Bradley, Michael, Gregg A. Jarrell and E Han Kim (1984), 'On the existence of an optimal capital structure: Theory and evidence', *Journal of Finance* **39**(3), 857–78.

Caine, Wesley K. (1981), 'Case for repealing the antidumping provisions of the tariff act of 1930', *Law and Policy in International Business* **3**.

Camerer, Colin (1999), *Uncertain Decisions: Bridging Theory and Experiments*, Kluwer Academic Publishers, chapter Ambiguity-Aversion and Non-Additive Probability: Experimental Evidence, Models and Applications, pp. 53–80.

Cover, Thomas M. and Joy A. Thomas (2006), *Elements of Information Theory*, Wiley Science.

Cremers, Martijn and Hongjun Yan (2009), 'Uncertainty and valuations', *Yale School of Management Working Papers* .

DeAngelo, Harry, Linda DeAngelo and Toni M. Whited (2011), 'Capital structure dynamics and transitory debt', *Journal of Financial Economics* **99**(2), 235–261.

Dempster, A. P., N. M. Laird and D. B. Rubin (1977), 'Maximum likelihood from incomplete data via the EM algorithm', *J. Roy. Statist. Soc. Ser. B* **39**(1), 1–38.

Dixit, Avinash (1992), 'Investment and hysteresis', *The Journal of Economic Perspectives* **6**(1), pp. 107–132.

Dixit, Avinish K. and Robert S. Pindyck (1994), *Investment Under Uncertainty*, Princeton University Press.

Duffie, Darrell (2001), *Dynamic Asset Pricing*, Princeton.

Duffie, Darrell and David Lando (2003), 'Term structures of credit spreads with incomplete accounting information', *Econometrica* **12**(4), 633–664.

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

Epstein, Larry G and Tan Wang (1994), 'Intertemporal asset pricing under knightian uncertainty', *Econometrica* **62**(2), 283–322.

Fama, Eugene F. and Kenneth R. French (2002), 'Testing trade-off and pecking order predictions about dividends and debt', *The Review of Financial Studies* **15**(1), pp. 1–33.

Faulkender, Michael and Mitchell A. Petersen (2006), 'Does the source of capital affect capital structure?', *Review of Financial Studies* **19**(1), 45–79.

Folland, Gerald B. (1999), *Real Analysis: Modern Techniques and Their Applications*, Wiley Inter-Science.

Frank, Murray Z. and Vidhan K. Goyal (2008), 'Trade-off and pecking order theories of debt', *Handbook of Corporate Finance* **2**, 911–922.

Garfinkel, Jon A. (2009), 'Measuring investors' opinion divergence', *Journal of Accounting Research* **47**(5), 1317–1348.

Gaspar, Jose-Miguel and Massimo Massa (2006), 'Idiosyncratic volatility and product market competition', *The Journal of Business* **79**(6), pp. 3125–3152.

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

Goldstein, Robert, Nengjiu Ju and Hayne Leland (2001), 'An ebit-based model of dynamic capital structure', *The Journal of Business* **74**(4), 483–512.

Graham, John (1996), 'Debt and the marginal tax rate', *Journal of Financial Economics* **41**(1), 41–73.

Graham, John R. (2000), 'How big are the tax benefits of debt?', *Journal of Finance* **55**(5), 1901–1941.

Greene, William H. (2012), *Econometric Analysis*, seventh edn, Peason.

Hamilton, James D. (1994), *Time Series Analysis*, Princeton University Press.

Hardt, John P. (1982), 'Soviet gas pipeline: U.S. options', *the library of congress congressional research service* .

Harvard Law Review (1983), 'Protecting steel: Time for a new approach', *Harvard Law Review* **96**(4), pp. 866–885.

He, Zhiguo, Si Li, Bin Wei and Jianfeng Yu (2010), 'Uncertainty, Risk, and Incentives: Theory and Evidence', *Booth School of Business, working paper* .

Hennessy, Christopher A. and Toni M. Whited (2005), 'Debt dynamics', *Journal of Finance* **60**(3), 1129–1165.

Hennessy, Christopher A. and Toni M. Whited (2007), 'How costly is external financing? evidence from a structural estimation', *Journal of Finance* **62**(4), 1705–1745.

Ju, Nengjiu and Jianjun Miao (2007), Ambiguity, learning, and asset returns, Boston University - Department of Economics - Working Papers Series wp2009-014.

Karatzas, Ioannis and Steven E. Shereve (1988), *Brownian Motion and Stochastic Caculus*, Springer-Verlag.

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

Knight, Frank (1921), *Risk, Uncertainty, and the Firm*, Dover Classics.

Kraus, Alan and Robert H. Litzenberger (1973), 'A state-preference model of optimal financial leverage', *The Journal of Finance* **28**(4), 911–922.

Leary, Mark T. and Michael R. Roberts (2005), 'Do firms rebalance their capital structures?', *Journal of Finance* **60**(6), 2575–2619.

Leland, Hayne E (1994), 'Corporate debt value, bond covenants, and optimal capital structure', *Journal of Finance* **49**(4), 1213–52.

Lemmon, Michael L. and Jaime F. Zender (2001), 'Looking under the lammpost: An empirical examination of the determinants of capital structure', *the David Eccles School of Busines, the University of Utah, Working paper series* .

Lemmon, Michael L., Michael R. Roberts and Jaime F. Zender (2008), 'Back to the beginning: Persistence and the cross-section of corporate capital structure', *Journal of Finance* **63**(4), 1575–1608.

Lowenfeld, Andreas F. (1983), *Public Controls on International Trade*, New York: M. Bender.

Markusen, Jame R. (1996), 'Comment: The wanning influence of big steel', *The Political Economy of Trade Policy* .

McDonald, Robert and Daniel Siegel (1986), 'The value of waiting to invest', *The Quarterly Journal of Economics* **101**(4), 707–27.

McLachlan, Geoff and Thriyambakam Krishnan (1997), *The EM Algorithm and Extensions*, John Wiley and Sons, Inc.

Miao, Jianjun and Neng Wang (2011), 'Risk, uncertainty, and option exercise', *Journal of Economic Dynamics and Control* **35**(4), 442–461.

Miller, Merton H. (1977), 'Debt and taxes', *The Journal of Finance* **32**(2), 261–275.

Minton, Bernadette A. and Karen H. Wruck (2001), 'Financial conservatism: Evidence on capital structure from low leverage firms', *AFA 2002 Atlanta; Dice Ctr. Working Paper* .

Moore, Michael (1996), 'The rise and fall of big steel's influence on u.s. trade policy', *The Political Economy of Trade Protection* pp. 15–34.

Morellec, Erwan (2004), 'Can managerial discretion explain observed leverage ratios?', *Review of Financial Studies* **17**(1), 257–294.

Nishimura, Kiyohiko G. and Hiroyuki Ozaki (2004), 'Search and knightian uncertainty', *Journal of Economic Theory* **119**(2), 299–333.

Nishimura, Kiyohiko G. and Hiroyuki Ozaki (2007), 'Irreversible investment and knightian uncertainty', *Journal of Economic Theory* **136**(1), 668–694.

Pastor, Lubos, Lucian A. Taylor and Pietro Veronesi (2009), 'Entrepreneurial learning, the ipo decision, and the post-ipo drop in firm profitability', *Review of Financial Studies* **22**(8), 3005–3046.

Petersen, Mitchell A. (2009), 'Estimating standard errors in finance panel data sets: Comparing approaches', *Review of Financial Studies* **22**(1), 435–480.

Pollard, David (2002), *User Guide to Measure Theoretic Probability*, Cambridge.

Rajan, Raghuram G and Luigi Zingales (1995), 'What do we know about capital structure? some evidence from international data', *Journal of Finance* **50**(5), 1421–60.

Range, John Jay (1980), 'Trigger price mechanism: Does it prevent dumping by foreign steelmakers ?', *North Carolina Journal of International Law and Commercial Regulation* **279**, 279–294.

Rigotti, Luca, Chris Shannon and Tomasz Strzalecki (2008), 'Subjective beliefs and ex ante trade', *Econometrica* **76**(5), 1167–1190.

Roberts, Michael R. and Tony M. Whited (2011), 'Endogenity in empirical corporate finance', *Mimeo* .

Rosenbaum, Paul and Donald Rubin (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**, 41–55.

Rubin, Donald B. and Neal Thomas (2000), 'Combining propensity score matching with additional adjustments for prognostic covariates.', *Journal of the American Statistical Association.* **95**, 573-585.

Solomon, Anthony M. (1978), 'Report to the president: A comprehensive program for the steel industry', *International Legal Materials* **17**(4), pp. 955–1015.

Strebulaev, Ilya A. (2007), 'Do tests of capital structure theory mean what they say?', *Journal of Finance* **62**(4), 1747–1787.

Strebulaev, Ilya A. and Baozhong Yang (2012), 'The mystery of zero-leverage firms', *Working Paper* .

Titman, Sheridan and Roberto Wessels (1988), 'The determinants of capital structure choice', *Journal of Finance* **43**(1), 1–19.

Trefler, Daniel (1993), 'Trade liberalization and the theory of endogenous protection: An econometric study of u.s. import policy', *Journal of Political Economy* **101**(1), 138–60.

Welch, Ivo (2004), 'Capital structure and stock returns', *Journal of Political Economy* **112**(1), 106–131.