ESSAYS IN CORPORATE FINANCE AND INVESTMENTS

BY

DONGMING SUN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Finance
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Associate Professor Dirk Hackbarth, Chair
Professor Heitor Almeida, Director of Research
Professor Neil Pearson
Assistant Professor Ola Bengtsson

Abstract

This dissertation studies the interplay of asset and liability sides of balance sheets, and considers both the level and the risk attributions of investments and financing sources.

The first chapter links financing frictions on the liability side to investment risk on the asset side. It studies the effect of financial constraints on equity holders' risk-shifting incentives within a real options framework. Within this framework, shareholders trade off the benefit of risk-shifting with the cost of financial constraints. Therefore risk-shifting is avoided ex post for highly constrained firms because the cost outweighs the benefit. In fact, both the risk-shifting incentive and the agency cost of risk-shifting are monotonically decreasing in financial constraints costs. In addition, the effect of debt maturity is also examined in this framework, and without financial constraints, there is no short-term debt effect. These model implications are supported in a large sample of firms over the 1965 to 2009 period: (1) financial constraints help to reduce risk-shifting incentives; (2) complementing the current view, financially unconstrained firms tend to shift risk even when they are still healthy; (3) short-term debt helps to strengthen the effect of financial constraints on reducing risk-shifting incentives; (4) the agency cost of risk-shifting is smaller for more constrained firms. The results are robust to the availability of internal financings.

The second chapter studies the opposite direction: the effect goes from the asset side to the liability side. It studies corporate investment and financing in a dynamic trade-off model with a sequence of irreversible investments. Conditional on future investment and financing opportunities, juvenile firms underutilize debt when financing investment the first time to retain financial flexibility. Underutilization of debt persists when adolescent firms mature (i.e. exercise their last investment options), and it is more (less) severe for more back-loaded (front-loaded) investment opportunities. Thus, leverage dynamics crucially hinge upon the structure of the investment process and otherwise identical firms appear to have significantly different target leverage ratios. Structural estimation of key parameters reveals that simulated model moments can match data moments. Furthermore, capital structure regressions using model simulated data based on these parameter estimates produce results in line with the empirical evidence, and explain the empirical puzzle that average leverage ratios are path dependent and persistent for very long periods of time.

The third chapter narrows down to the liability side and studies the puzzle of whether idiosyncratic risk predicts the cross-section of stock returns and the direction of the prediction. This chapter examines this relationship by extracting implied idiosyncratic variances from option prices and decomposing past realized idiosyncratic variances into expected and unexpected components. The Fama-MacBeth (1973) regressions using different samples show mixed results. The significant positive (negative) relationship between cross-sectional stock returns and implied idiosyncratic variance (past realized idiosyncratic variance) is mainly driven by the sample with low (high) idiosyncratic variances. It is plausible that the mixed results in the literature are caused by the conflicting effects

of implied idiosyncratic variance for low idiosyncratic variance stocks and persistent idiosyncratic variance shock for high idiosyncratic variance stocks.

To Father and Mother.

Acknowledgments

This thesis would not have been possible without the support of many people. I am deeply grateful to my adviser, Dr. Dirk Hackbarth, for his continuing guidance, encouragement, and support. His vision and enthusiasm in scientific research inspired and motivated me throughout my pursuit of the degree. Also thanks to the other committee members, Neil Pearson, Heitor Almeida, and Ola Bengtsson, who offered guidance and invaluable insights in many areas. I also want to thank the other faculty members in the Department of Finance at UIUC for their suggestions and support. The staff in the Department of Finance has always been extremely helpful and friendly. I am especially thankful to Andrea Kelly, Karen Brunner, Shelley Campbell and Sharon Lee Allen. Thanks to the Department of Finance providing me with the financial means to complete this thesis.

During the years in Illinois, I have been extremely lucky to have many supportive and helpful friends. I would like to especially thank Chunfang Lin, Yue Zhuo, Bei Ye, Bruno Laranjeira, Yue Chen, Lu Wang and Lina Deng who went out of their way to help me going through tough times. And finally, thanks to my parents, Xiujie Zhang and Fenglin Sun, and my boy friend, Jiading Gai, for their great love, unselfish support, endless patience, and strong encouragement through my life. I am forever indebted to them and would dedicate all my achievements to them.

Table of Contents

Chapter 1

Financial Constraints, Risk-shifting and Debt Maturity

1.1 Introduction

The interplay of firms' financing and investment decisions has been a major interest in modern corporate finance. Researchers have done considerable studies to gauge the effect of investment decisions on optimal capital structure and generated interesting predictions. One direction of such research focuses on the risk aspect of investments, i.e. the effect of equity holders' risk-shifting preferences (asset substitution) on optimal capital structure decisions. However, the opposite direction, namely the effect of financing activities on investment risk decisions has not been fully explored. This paper provides a link in this direction and quantifies the effect of financial constraints on reducing risk-shifting incentives as well as its interaction with debt maturity effect. This link produces new perspectives to look at risk-shifting problem.

Risk choice of corporate investments is a key element in the risk-shifting (asset substitution) literature. Galai and Masulis (1976) and Jensen and Meckling (1976) pose the risk-shifting problem. They suggest that the payoff structure of different claim holders induces equity holders of a levered firm to extract value from current debt holders; they do so by making risky investments in order to catch the upside benefits and leave the downside risk to debt holders when the firm is near bankruptcy. This behavior of equity holders is detrimental to firm value and creates an agency cost. Therefore it is essential to evaluate the magnitude of this agency cost and find ways to reduce/remove it.

Some studies have sought to identify factors that help to mitigate this problem, e.g. bond covenants in Smith and Warner (1979), convertible debt in Hennessy and Tserlukevich (2004), debt maturity in Barnea, Haugen, and Senbet (1980) and Barclay and Smith (1995a), and managerial compensation in Subramanian (2003). Surprisingly, none of the studies has explored the risk-shifting problem in the perspective of financing frictions. This paper makes the first attempt to link financial constraints, one type of financing frictions, to the risk-shifting literature. It formally derives and empirically tests the effect of financial constraints on corporate risk-shifting behaviors. It also examines the interaction of this effect with debt maturity effect. The theoretical part formulates this problem within a real options framework, which allows for a simple model setup to incorporate financial constraints, debt maturity and risk-shifting incentives. Furthermore, it generates closed-form solutions for debt, equity and firm values. Based on these values, the optimal risk-shifting and default thresholds can be solved simultaneously with simple conditions.

Within this framework, the degree of financial constraints is represented by an exogenous cost of external financing M following Hennessy and Whited (2007). This representation encompasses not only the "cost constraint", i.e. constrained firms have to pay a dead-weight cost beyond the fair cost of external financing, but also covers the "quantity constraint" where M is extremely high, i.e.

firms cannot raise external funds beyond a certain point no matter how much they are willing to pay for it. The firm is unconstrained if M is extremely low. Almeida and Campello (2005) provide discussions of the "cost constraint" v.s. the "quantity constraint" in more detail. The inclusion of M complements the existing literature and generates new predictions for the role that financial constraints play in the interplay of financing and investments.

Risk-shifting, the other key component of this study, is defined as an irreversible risk-increasing investment opportunity in the model. It is called risk-shifting option throughout the paper. In the model the firm operates at a low risk level initially and the equity holders determine the initial coupon of a perpetual debt by optimizing the initial firm value. It is assumed that the ex post risk-shifting behaviors cannot be credibly excluded through contracting with debt holders ex ante. So once debt is in place, equity holders can choose to exercise the risk-shifting option and raise the firm's risk level permanently by investing in a risky project fully funded with equity.

The above two key components, financial constraints and risk-shifting option, yield a trade-off between the benefit of exercising the risk-shifting option and the cost of financial constraints M. Given the similarity between equity and option, the equity holders profit from exercising the risk-shifting option at the cost of M. When firm is less constrained, the gain from risk-shifting outweighs the cost; equity holders exercise the risk-shifting option when X falls below the optimal risk-shifting threshold XR for the first time. In contrast, when the cost M dominates the benefit, equity holders abandon the risk-shifting option. This trade-off provides a channel through which financial constraints help to reduce the risk-shifting incentives ex post.

In fact, this trade-off is studied with the presence of debt maturity in the model. Debt maturity is introduced through an exogenously given debt rollover rate m by following the approach in Leland (1994b). Once the optimal initial debt with par value P is in place, a portion m of P is retired and reissued continuously with the same coupon and principal amount. This mechanism implies an average debt maturity of $T \approx 1/m$ years. Three rollover rates m are used in the numerical example to represent long-term, medium-term and short-term debt maturities in this paper. With this mechanism, shorter debt maturity entails higher periodic payments from equity holders, which can potentially reduce the benefit of risk-shifting.

Combining all three factors above, this model reveals interesting implications of the relationship between financial constraints, debt maturity and risk-shifting incentives:

1. Risk-shifting incentive decreases with financial constraints. That is, financially constrained firms are less likely to engage in risk-shifting behaviors than financially unconstrained firms.

2. Financially unconstrained firms tend to risk-shift even when they are healthy. So financial distress is not the sufficient condition for risk-shifting.

3. Without the presence of financial constraints, the short-term debt effect in Barnea, Haugen and Senbet(1980) does not exist. Only when financial constraints are controlled for, short-term debt complements financial constraints in reducing risk-shifting incentives.

4. The agency cost of risk-shifting is lower for financially constrained firms. They have lower loss in optimal leverages and firm values compared to financially unconstrained firms.

To test these model implications empirically, I use a large sample of firms from the intersection of COMPUSTAT annual tape and CRSP for the period between 1965 and 2009 and classify them into financially constrained and unconstrained firms based on four financial constraints criteria. The effect of financial constraints on risk-shifting is studied in the investment-expected volatility regression framework in Eisdorfer (2008). The interaction of expected volatility and financial constraints index is significantly negative for financially constrained firms implying that the risk-shifting incentive is low for such firms. Furthermore, when financial constraints are controlled for, financial distress as indicated by Altman's Z-score loses its explanatory power in the regressions. What's more, the debt maturity effect is introduced by adding dummy variables into the regressions to represent weighted-average maturities of one, two, three, four and five years. As predicted by the model, the unconditional short-term debt maturity effect does not exist. It shows up only when financial constraints are controlled for. In addition, I study the agency cost of risk-shifting by running regressions of leverage and asset returns on contemporaneous investment with control of financial constraints and other variables. Consistent with the predictions of the model, the empirical tests also reveal that financially constrained firms have 0.152% higher of leverage and 0.605% higher of asset return per 1% investment the firm makes than unconstrained firms, which can be a big reduction in agency costs if the firm makes heavy investments.

The model of this study imposes a strong assumption that equity holders rely on external financing to realize risk-shifting activities, so concerns might arise that these results do not apply to firms with high internal funds. However, the descriptive statistics show that financially constrained firms have more cash holdings on average and this is exactly the precautionary step they take due to financial constraints. Financially unconstrained firms, on the other hand, have less incentive to indulge their cash holdings because the cost of external financing is relatively low for them. To test this claim empirically, I control for cash holdings in the empirical tests and find that the effect of financial constraints still hold even when cash holding is controlled for.

These findings illustrate a novel link between the financial constraints literature and the risk-shifting literature. The financial constraints literature mainly focuses on the level of investments and proves the dark side of financial constraints. For example, Campello, Graham and Harvey (2010) use experimental design to provide direct evidence of the effect of financial constraints on real investments by surveying 1,050 CFOs in the US, Europe and Asia during the latest credit crunch. They find that constrained firms experience higher cut in capital spending and pass on attractive investment opportunities during the crisis. Another line of financial constraints research shows that financial constraints restrain investment by studying the investment-cash flow sensitivity. Fazzari, Hubbard and Petersen (FHP) (1988) propose that investment decisions within financially constrained firms should be influenced not only by the virtue of investment opportunities but also by the availability of internal funds. There is a monotonic relation between the degree of financial constraints and the investment-cash flow sensitivity. Thus the effect of financial constraints on investment spending can be examined by comparing investment-cash flow sensitivities across sub-samples sorted on characteristics related to financial constraints. However, the monotonic relation of financial constraints and investment-cash flow sensitivity is challenged by Kaplan and Zingales (1997) among others (e.g. Erickson and Whited (2000), Gomes(2001) and Alti (2003)). By carrying out an in-depth investigation of the high investment-cash flow sensitivity sub-sample of that in FHP (1988), they find contradictory results, casting a doubt on the robustness of FHP (1988). To get around the Kaplan and Zingales (1997) critique, Almeida, Campello and Weisbach (2004) propose the cash flow sensitivity of cash to capture the effect of financial constraints. A more recent work by Almeida and Campello (2007) suggests that asset tangibility has a differential effect on investment-cash flow sensitivities because asset tangibility increases a firm's ability to get external finance when it is financially constrained but has no effect when it is unconstrained. Campello and Hackbarth (2008) formulate the *credit multiplier* effect induced by asset tangibility within a real options framework and show that investment spending is improved with asset tangibility for financially constrained firms.

These studies, however, overlook the risk profile of investments. When a firm chooses a very risky investment, it changes the riskiness of the firm as well. Its impact on the valuations of equity and debt shows the agency cost to the firm. This attribution of investment can be discovered in the risk-shifting literature.

Some theoretical studies in the risk-shifting literature try to quantify the agency cost of risk-shifting, but find mixed results. Leland (1998) investigates it by studying the effect of risk-shifting (asset substitution) on optimal capital structure, debt maturity and risk management choice using a real options framework. He introduces a reversible risk-shifting option into a dynamic framework in which bonds are callable, and finds that risk-shifting restricts leverage and debt maturity and increases yield spreads, but the effects are small. Ericsson (2000) also studies the simultaneous choice of leverage and debt maturity in the presence of risk-shifting incentives in a real options framework. However, the risk-shifting option in his model is irreversible and bonds can not be called. Contrary to Leland (1998) he finds that, by ruling out risk-shifting, firms can take 20% more leverage and use distinctively longer-term debt financings. In both models, risk-shifting is free. This paper differs from the above two in that risk-shifting option is subject to a financial constraints cost M when exercised. Moreover, risk-shifting is irreversible, the decision to default is endogenous, and the criterion for abandoning the risk-shifting option is different.

On the other hand, Eisdorfer (2008) presents empirical evidence of the existence of risk-shifting behaviors in financially distressed firms and estimates the cost of risk-shifting to debt holders to be 6.38% during high volatility periods. He argues that without risk-shifting, the investment-volatility relation is negative because the option value of waiting-to-invest dominates when volatility is high; however, when risk-shifting benefit presents, the relation can be reversed. This provides an ideal framework for the empirical tests in this paper.

The remainder of the paper is organized as follows. Section 3.2 presents the model setup, derives the optimal risk-shifting and default thresholds, and discusses the model implications. Section 1.3 uses a large dataset to test the hypotheses empirically. Section 1.4 presents some robustness tests and section 1.5 concludes.

1.2 Model

In this section I consider two real options models: a simple benchmark model with no risk-shifting prospect, and a risk-shifting model with exogenously given financial constraints costs M. The real options framework provides a natural venue to tackle the linkage between financial constraint, debt maturity and risk-shifting behaviors. It abstracts equity holders' choice of risk as an option to increase risk and generates closed-form solutions of debt and equity values.

Following the EBIT-based model in Goldstein, Ju and Leland (2001), a typical firm has cash flows (EBIT) generated by its assets evolving stochastically according to a log-normal process:

$$\frac{dX(t)}{X(t)} = \mu \; dt + \sigma \; dW(t), \tag{1.1}$$

where μ is the expected rate of return, σ is the risk of return, and $dW(t)$ is the increment of a standard Wiener process.

At time 0, the firm has an initial cash flow X_0 and a low risk level σ_L. It faces a marginal tax rate τ and a risk-free rate r_f. The firm is levered and it uses a debt with perpetual coupon payment C and a principal P, a fraction m of which is retired and reissued continuously with the same structure. The optimal coupon rate C is determined by maximizing the initial firm value for the exogenously given debt rollover rate m. Once the initial debt is in place, equity holders can choose to default endogenously at an optimal ex post default threshold XD and the percentage cost of bankruptcy is α. The firm is also subject to an exogenous financial constraints cost M for future investments.

1.2.1 The Benchmark Model

To quantify the effect of financial constraints on risk-shifting behavior and its interaction with debt maturity, I start with a simple benchmark model. In this model, equity holders can credibly pledge against risk-shifting behaviors ex ante. Their only decision ex post is to determine when to default[1], i.e. what the optimal endogenous default threshold XD is. To do so, equity and debt values need to be derived first.

The Valuations

According to Goldstein et.al. (2001), equity, debt and firm values can be solved with the following Ordinary Differential Equation (ODE):

$$\mu X V_X + \frac{\sigma^2}{2} X^2 V_{XX} - rV + \Pi = 0, \tag{1.2}$$

where V can be the equity value SB, the debt value DB or the firm value VB(B stands for benchmark model), and Π denotes the intermediate cash flows to the claimants. Due to the debt rollover feature of the model, the current debt holders receive not only the periodic coupon payments ($\Pi = C$), but also the repayment of a fraction m of the principal mP whereas the equity holders claim on the residual after-tax cash flows.

Debt Value

Following Leland (1998), the general solution for debt value $DB(X)$ can be written as:

$$DB(X, C, m) = A_1 X^{a_m} + A_2 X^{z_m} + \frac{C + mP}{r + m}, \qquad X > XD, \tag{1.3}$$

where

$$a_m = \frac{-(r - \frac{\sigma^2}{2}) - \sqrt{(r - \frac{\sigma^2}{2}) + 2\sigma(r + m)}}{\sigma^2}, \tag{1.4}$$

$$z_m = \frac{-(r - \frac{\sigma^2}{2}) + \sqrt{(r - \frac{\sigma^2}{2}) + 2\sigma(r + m)}}{\sigma^2} \tag{1.5}$$

[1]To simplify the model and focus the analysis on risk-shifting, I do not include any investment options other than the one for risk-shifting in the model.

and the $\frac{C+mP}{r+m}$ portion stands for the present value of the intermediate cash flows to debt holders. The constants A_1 and A_2 can be computed with the boundary conditions when $X = XD$ and $X \uparrow \infty$.

When the firm is highly profitable ($X \uparrow \infty$), its debt can be viewed as a perpetuity with continuous payments of $C + mP$. On the other hand, when the firm is in bankruptcy ($X = XD$), debt holders seize the liquidation value of the firm after paying off the bankruptcy cost, and the equity holders are left with nothing. These two boundary conditions can be summarized as follows:

$$DB(\infty, C, m) = \frac{C + mP}{r + m}, \tag{1.6}$$

$$DB(XD, C, m) = (1 - \alpha)(1 - \tau)\frac{XD}{r - \mu}. \tag{1.7}$$

Therefore,

$$DB(X, C, m) = \frac{C + mP}{r + m} + (\frac{X}{XD})^{a_m}\Big[(1 - \alpha)(1 - \tau)\frac{XD}{r - \mu} - \frac{C + mP}{r + m}\Big], \ X \geq XD, \tag{1.8}$$

where

$$\begin{aligned} P &= DB(X_0, C, m) \\ &= \frac{\frac{C}{r+m}\Big[1 - (\frac{X_0}{XD})^{a_m}\Big] + (\frac{X_0}{XD})^{a_m}(1 - \alpha)(1 - \tau)\frac{XD}{r-\mu}}{\frac{r}{r+m} + \frac{m}{r+m}(\frac{X_0}{XD})^{a_m}}. \end{aligned} \tag{1.9}$$

The debt value shown in equation (1.8) contains three parts: a risk-free bond value, the contingent claim of the liquidation value excluding the bankruptcy cost and the loss of future coupons contingent on default.

Firm Value

Similar to debt value, the general solution to firm value is as follows:

$$VB(X, C, m) = K_1X^a + K_2X^z + \frac{(1 - \tau)X}{r - \mu}, \ X \geq XD, \tag{1.10}$$

where

$$a = \frac{-(r - \frac{\sigma^2}{2}) - \sqrt{(r - \frac{\sigma^2}{2}) + 2\sigma r}}{\sigma^2}, \tag{1.11}$$

$$z = \frac{-(r - \frac{\sigma^2}{2}) + \sqrt{(r - \frac{\sigma^2}{2}) + 2\sigma r}}{\sigma^2} \tag{1.12}$$

and $\frac{(1-\tau)X}{r-\mu}$ is the present value of the intermediate cash flows to the firm. The two coefficients K_1 and K_2 can be solved by the following two boundary conditions:

$$VB(\infty, C, m) = \frac{(1 - \tau)X}{r - \mu}, \tag{1.13}$$

$$VB(XD, C, m) = (1 - \alpha)(1 - \tau)\frac{XD}{r - \mu}. \tag{1.14}$$

The firm value can thus be expressed as:

$$VB(X, C, m) = \frac{(1-\tau)X}{r-\mu} + \frac{\tau C}{r} - (\frac{X}{XD})^a \Big[\frac{\alpha(1-\tau)XD}{r-\mu} + \frac{\tau C}{r}\Big], \; X \geq XD. \tag{1.15}$$

As shown above, the pre-default firm value can be decomposed into three parts: the unlevered firm value $\frac{(1-\tau)X}{r-\mu}$, the tax benefit of debt $\frac{\tau C}{r}$, and the contingent loss at default including the bankruptcy cost and the loss of tax benefit of debt. At default, the firm value is simply the after-tax liquidation value minus the bankruptcy cost.

The equity value is the difference between the firm value and the debt value, i.e.

$$SB(X, C, m) = VB(X, C, m) - DB(X, C, m). \tag{1.16}$$

Due to the simplification of the model, the valuations above are independent of the financial constraints cost M because no external financing is needed once debt is in place in the benchmark model.

The Optimization Problem

The objective of the firm is to get the first-best valuation by choosing the optimal capital structure ex ante, i.e. the coupon C^* for the given debt rollover rate m, subject to the condition that the default threshold XD is determined optimally ex post. This problem can be formulated as follows:

$$\begin{aligned} C^* &= \max_C VB(X_0, C, m) \\ s.t. & \quad \frac{\partial SB(X, C, m)}{\partial X}\Big|_{X=XD} = 0 \end{aligned} \tag{1.17}$$

The ex post optimal default threshold XD must satisfy condition (1.17) for any level of C. When X is higher than XD, the future cash flows are still higher than the coupon payments and the bankruptcy costs. Therefore it is worthwhile for the equity holders to wait and default at a lower threshold. On the contrary, if they hold the firm for too long ($X < XD$), the gains of keeping the firm alive is outweighed by the coupon payments and the bankruptcy costs. They should exercise the default option earlier. As a result, the equity value is maximized at XD for a given C and m. Here's the formula for XD as a function of C and m:

$$XD(C, m) = \frac{a\frac{\tau C}{r} - a_m \frac{C+mP}{r+m}}{\frac{1-\tau}{r-\mu}\Big[1 - a\alpha - a_m(1-\alpha)\Big]}. \tag{1.18}$$

1.2.2 The Risk-shifting Model

This section continues the discussion by adding the cost of financial constraints M and a risk-shifting option to the benchmark model. This model gives more insights into the strength of risk-shifting incentives and the magnitude of the agency cost of risk-shifting in the presence of financial constraints and debt maturity. I will describe the model by explaining how financial constraints, risk-shifting and debt maturity are modeled first.

The financial constraints cost M

In an imperfect capital market, firms facing financial constraints are subject to significant costs for external financings for reasons such as lack of credit-worthiness, low asset tangibility and bad economic environment (e.g. Hennessy and Whited (2007)). As a result, financially constrained firms have to pay the added cost beyond the fair external financing costs. This cost is modeled as an exogenously given constant M which represents the degree of financial constraints. M is larger for more constrained firms and lower for unconstrained firms.

The definition of M encompasses not only the "cost constraint" discussed above, but also the "quantity constraint" discussed in Almeida and Campello (2005), i.e. firms' inability to raise external funds regardless of how much cost they are willing to bear. In fact the "quantity constraint" is the extreme case of $M \uparrow \infty$. The other extreme case, $M \downarrow 0$, gives the unconstrained situation.

The Risk-shifting Option

Considering a firm with an initial perpetual debt having coupon C, the debt holders cannot prevent the risk-shifting behaviors ex ante by contracting with shareholders, or such prevention can not be credibly enforced ex post. Under this assumption, the equity holders have a risk-shifting option to increase investment risk from σ_L to σ_H by making a risky investment after the initial debt is in place. This risk-shifting investment is fully financed by equity and the equity holders bear the financial constraints cost M [2] if the option is exercised. The ex post optimal risk-shifting and default thresholds are determined endogenously by shareholders to maximize the equity value. Following Ericsson (2000), I also assume that risk-shifting is irreversible: once the risk is increased, it will stay at that level as long as the firm is alive. The addition of this option creates a direct contrast to the benchmark model. The differences between the two models reveal the real impact of risk-shifting behaviors.

The Mechanism at Work

In the benchmark model, the equity holders' only concern is when to default. An additional decision is in place in the risk-shifting model: shareholders may decide to switch the investment risk to a higher level before they default on debt. This action makes both the equity[3] and the default option more valuable. As a result, the equity holders hold on to the default option longer, i.e. the post-shifting default threshold is lower than the benchmark case. This entitles the equity holders to more cash flows and in turn increases the value of the equity. However this comes with a price, i.e. the financial constraints cost M. When M is affordable, the gains from risk-shifting outweighs M and the equity holders exercise the risk-shifting option when X falls to the optimal risk-shifting threshold XR. In contrast, when M is high so as to dominate the benefits of risk-shifting, the equity holders will abandon the risk-shifting option. This trade-off provides a channel through which financial constraints helps to reduce the risk-shifting incentives ex post.

[2]Note that it is reasonable to assume that the equity holders bear all the financial constraints cost M for risk-shifting. When risk-shifting behaviors cannot be credibly excluded ex post, bond holders will take it into account and request higher return ex ante. This leaves all costs to the equity holders.

[3]Based on Merton (1973), equity can be modeled as an option on firm value. The value of this option increases with volatility.

Debt Maturity

The current literature on risk-shifting suggests that short-term debt helps to reduce risk-shifting incentives (i.e., Barnea, Haugen, and Senbet (1980)), but it receives mixed empirical support. To study how debt maturity interacts with financial constraints, I introduced a continuous debt rollover rate m as in Leland (1994b). In this framework, optimal coupon C is determined initially by maximizing the total firm-value. The debt principal is always kept at its par value P. At any instant, a fraction m of the principal is retired at par and reissued at market value with identical characteristics, thus the average debt maturity is $T \approx \frac{1}{m}$. The market value can be different from the par value, and the difference accrues to equity holders if this is the case. Since this paper does not attempt to study the optimal debt maturity, the refinancing cost and debt rollover cost are ignored in order to get a clean view on the mechanism at work. Should they have been introduced, they would have made short-term debt more costly to use and worked against the short-term debt effect.

Valuations

Within the same contingent claim valuation framework as in section (1.2.1), the equity, debt and firm values in this model satisfy the ODE (1.2) and can be expressed in the same general form of equation (1.3), but follow different boundary conditions in. I use the suffix "H" to denote the valuations for post-shifting (high volatility) environment and use "L" for valuations before the risk-shifting option is exercised (low volatility). The option exercising is triggered if X decreases to XR from above for the first time. After that the firm operates in the benchmark model environment but with the higher risk level σ_H because there is no risk-shifting options any more. The optimal post-shifting default threshold is denoted by XDH. The implicit assumption here is $XR > XDH$. That is, risk-shifting can only happen before the equity holders decide to default. When $XR \leq XDH$, it is optimal not to shift risk at all and the firm operates in the same fashion as in the benchmark environment.

After Risk-shifting

The risk-shifting model can be solved using backward induction. After risk-shifting, the equity holders are left with the default option only. This is the same case as the benchmark model, except that the risk is higher, i.e. σ_H, and the default threshold XDH is lower than XD. Thus the post-shifting debt, equity and firm values, i.e. $DH(X, C, m)$, $SH(X, C, m)$ and $VH(X, C, m)$, take the same functional forms as in the benchmark case except that σ should be replaced by σ_H.

Before Risk-shifting

When the firm's profitability is well above the risk-shifting threshold XR, it is not optimal for the shareholders to take high risk because the gain from risk-shifting is outweighed by the higher bankruptcy cost and the loss of future profits. In this case, the equity holders have both the risk-shifting option and the default option in hand. When $X \uparrow \infty$, the firm is risk-free and the total firm value is simply the present value of the payout streams to equity and debt holders, i.e. $VL(X \uparrow \infty, C, m) = \frac{(1-\tau)X}{r-\mu}$. To the other extreme when $X = XR$, equity holders will exercise the risk-shifting option. At this point the firm value should transit smoothly in the vicinity of XR to avoid arbitrage opportunities. Hence it must satisfy the value-matching condition $VL(XR, C, m) = VH(XR, C, m) - M$. Similarly, the debt value should equal to the risk-free bond price $\frac{C+mP}{r+m}$ as $X \uparrow \infty$ and it should satisfy the value-matching condition at XR: $DL(XR, C, m) = DH(XR, C, m)$.

With the above mentioned conditions and the general solution (1.3), I obtain the closed-form solutions of debt and firm values for the pre-and post-shifting stages:

$$
\begin{aligned}
DL(X,C,m) &= \frac{C+mP}{r+m} + \Phi_m\Big[\frac{(1-\alpha)(1-\tau)XDH}{r-\mu} - \frac{C+mP}{r+m}\Big],\ X \leq XR, \\
&\qquad (1.19) \\
DH(X,C,m) &= \frac{C+mP}{r+m} + (\frac{X}{XDH})^{a_{Hm}}\Big[\frac{(1-\alpha)(1-\tau)XDH}{r-\mu} - \frac{C+mP}{r+m}\Big],\ X \leq XDH, \qquad (1.20) \\
VL(X,C,m) &= \frac{(1-\tau)X}{r-\mu} + \frac{\tau C}{r}\Big[1-\Phi\Big] - \frac{\alpha(1-\tau)XDH}{r-\mu}\Phi - (\frac{X}{XR})^{a_L}M,\ X \leq XR, \qquad (1.21) \\
VH(X,C,m) &= \frac{(1-\tau)X}{r-\mu} + \frac{\tau C}{r}\Big[1-(\frac{X}{XDH})^{a_H}\Big] - \frac{\alpha(1-\tau)XDH}{r-\mu}(\frac{X}{XDH})^{a_H},\ X \leq XDH, \qquad (1.22) \\
&\qquad (1.23)
\end{aligned}
$$

where $\Phi = (\frac{X}{XR})^{a_L}(\frac{XR}{XDH})^{a_H}$ and $\Phi_m = (\frac{X}{XR})^{a_{Lm}}(\frac{XR}{XDH})^{a_{Hm}}$ are the hitting claims which equal to 1 if equity holders shift risk at XR and zero otherwise. The major difference between Φ and Φ_m lies in whether the debt rollover rate m is factored in the definitions of a and z. The principal of debt is fixed at the initial debt value, i.e.

$$
\begin{aligned}
P &= DL(X_0,C,m) \\
&= \frac{(1-\Phi_m)\frac{C}{r+m} + \Phi_m\frac{(1-\alpha)(1-\tau)XDH}{r-\mu}}{\frac{r}{r+m} + \frac{m}{r+m}\Phi_m}. \qquad (1.24)
\end{aligned}
$$

The equity values are simply the difference between the firm and debt values: $SL(X,C,m) = VL(X,C,m) - DL(X,C,m)$ and $SH(X,C,m) = VH(X,C,m) - DH(X,C,m)$.

As shown in equation (1.21) the firm value is composed of four parts: the unlevered firm value, the tax benefit of debt, the default cost and the financial constraints cost contingent on risk-shifting. Once the risk-shifting option is exercised, the firm value contains only the first three components. The debt values in both stages are composed of three parts: the risk-free bond price, the loss of future coupons contingent on bankruptcy and the liquidation value seized by the debt holders at default excluding bankruptcy costs. They differ only in the hitting claims. In the pre-shifting stage, the hitting claim contains the option of risk-shifting, whereas in the post-shifting stage, the hitting claim is related to the default-option only.

The Optimization Problem

The closed-form solutions of the equity, debt and firm values in the last subsection make it possible to quantify the optimal coupon and the default and risk-shifting thresholds. This can be achieved by solving the following optimization problem:

$$
\begin{aligned}
&\max_C VL(X_0,C,m) \\
s.t.\quad &\frac{\partial SH}{\partial X}\Big|_{X=XDH} = 0, \qquad (1.25) \\
&\frac{\partial SL}{\partial X}\Big|_{X=XR} = \frac{\partial SH}{\partial X}\Big|_{X=XR}. \qquad (1.26)
\end{aligned}
$$

The first condition (1.25) shows that the equity holders choose to default in the post-shifting stage when the equity value is maximized. Please refer to the subsection (1.2.1) for the reasoning.

This condition solves the optimal post-shifting default threshold:

$$XDH(C,m) = \frac{a_H \frac{\tau C}{r} - a_{Hm} \frac{C+mP}{r+m}}{\frac{1-\tau}{r-\mu}\Big[1 - \alpha a_H - (1-\alpha) a_{Hm}\Big]}. \tag{1.27}$$

If the profitability deteriorates after the firm becomes more risky, the equity holders optimally stop fulfilling their debt service at XDH. From equation (1.27) it is easy to see that XDH is linear in coupon C and independent of M. Hence highly levered firms are more likely to default ex post.

The second condition (1.26) solves the optimal risk-shifting threshold. It follows from the value-matching condition to make sure that XR is time-invariant(Dixit (1993)). At this point the benefit of expropriating values from debt balances off the financial constraints cost M and the increased loss from bankruptcy. This condition gives the following equation that XR must follow:

$$\begin{aligned} & \Big[\frac{\tau C}{r} + \frac{\alpha(1-\tau)XDH}{r-\mu}\Big]\Big(\frac{XR}{XDH}\Big)^{a_H}(a_H - a_L) - a_L M \\ = & \Big[\frac{(1-\alpha)(1-\tau)XDH}{r-\mu} - \frac{C+mP}{r+m}\Big]\Big(\frac{XR}{XDH}\Big)^{a_{Hm}}(a_{Lm} - a_{Hm}). \end{aligned} \tag{1.28}$$

Notice that both XDH and XR are functions of the ex ante coupon C. Each C produces a set of thresholds and values. The optimal ex ante coupon C can be determined by varying C and comparing the corresponding firm values $VL(X_0, C, m)$.

The decision of Risk-shifting

The previous subsection presents the condition for determining the optimal risk-shifting threshold XR. If $XR \leq XDH$, it will never be optimal for the equity holders to shift risk. On the other hand when $XR > XDH$, if equity holders cannot credibly pledge against risk-shifting behavior ex ante, debt holders will assume that equity holders implement the risk-shifting model: as soon as X decreases to XR for the first time, equity holders will make the risk-increasing investment. As a result, the firm can only issue $P = DL(X_0, C, m)$ where C is determined by the optimization problem (1.25).

However, the condition $XR > XDH$ does not guarantee risk-shifting because it is only an option. In fact, if equity holders receive higher equity value by abandoning the risk-shifting option, they will implement the benchmark model ex post at the debt level determined by the risk-shifting model. The optimal default threshold in this case is denoted by XDL and determined by maximizing the pre-shifting equity value:

$$\frac{\partial SB(a_L)}{\partial X}\bigg|_{X=XDL} = 0. \tag{1.29}$$

Note that the firm mimics the benchmark model using the optimal coupon C from the risk-shifting model and setting $\sigma = \sigma_L$ in equation (1.18). That is,

$$XDL(C) = \frac{a \frac{\tau C}{r} - a_{Lm} \frac{C+mP}{r+m}}{\frac{1-\tau}{r-\mu}\Big[1 - a\alpha - a_{Lm}(1-\alpha)\Big]}. \tag{1.30}$$

This is only a sub-optimal solution compared to the base model, but it is still possible that for some $X > XDL$ and $XR > XDH$, $SB(X_0, C, m) > SL(X_0, C, m)$ and $VB(X_0, C, m) > VL(X_0, C, m)$. That is, even a sub-optimal solution of the base model may dominate the risk-shifting model. When this happens, the equity holders will abandon the risk-shifting option.

Based on the above discussions we can see that XR can be used as a measure of risk-shifting incentive. The higher XR is, the more likely that the equity holders will engage in risk-shifting ex post. In addition, the relative measure $K = \frac{XR}{XDL}$ depicts how far risk-shifting is from default. This gives us information on how likely equity holders will shift risk when the firm is relatively healthy.

The Magnitude of the Agency Cost of Risk-shifting

The next effect to examine is the magnitude of the agency cost of risk-shifting. According to the literature (e.g. Leland (1998), Ericsson (2000) and Eisdorfer (2008)), shareholders' risk-shifting behavior transfers value from bondholders to equity holders, and destroys firm value at the same time. Therefore the agency cost can be measured by the reduction of values because of risk-shifting. That is, the difference of debt or firm values between the risk-shifting model and the benchmark model quantifies the agency cost of risk-shifting. Equities, on the other hand, benefit from this behavior thus have a value appreciation. To combine all three effects in a unified measure, it is appropriate to use leverage, the ratio of debt value to firm value (debt plus equity), to specify the cost. More specifically, the reduction of the optimal ex ante leverage indicates the cost of risk-shifting. In fact, the effect of the financial constraints cost M is implicit in the valuations of the cost.

Due to the complexity of the algebraic expressions, I will illustrate the cost of risk-shifting with a numerical example in the next subsection.

1.2.3 A Numerical Example

Consider a representative firm with initial cash flow normalized to $X_0 = 100$. The firm faces a marginal tax rate $\tau = 20\%$ and a percentage bankruptcy cost $\alpha = 0.25$. The expected rate of return on its asset is $\mu = 1\%$, and the risk-free rate is $r_f = 6\%$. The firm has an option to switch its current low risk level $\sigma_L = 25\%$ to a higher one $\sigma_H = 40\%$. To make it more realistic, the financial constraints cost imposed on the firm is expressed as $M = M_0(1 + i)$, where $i \in [0, 0.6]$. In this example, the risk-shifting cost M for the relatively unconstrained firm is $M_0 = 30$, and depending on the degree of financial constraints it can be up to 60% more costly to do risk-shifting.

Figure 1.1 displays the optimal firm value for three debt maturities represented by $m = 0, 0.1$ and 0.2 across a range of financial constraints costs. Firm value goes up as financial constraints cost goes up, and there is a big jump in value when equity holders switch to the base model. The bottom panel shows exactly when the equity holders give up risk-shifting for the ten-year debt ($m = 0.1$) case: they switch to the benchmark model as soon as the equity value of the risk-shifting model $SL(C)$ drops below the equity value of the base model $SB(C)$. From the figure, we can see that the firm using the short-term debt ($m = 0.2$) switches at a lower financial constraints cost ($i = 15\%$), the medium debt at the highest financial constraints cost ($i = 55\%$) and the long-term debt in the middle. Therefore, there is no monotonic relation between risk-shifting incentive and debt maturity. The unconditional short-term debt effect does not hold in this model. The dotted line in the top panel represents the optimal firm value in the base model where there is no risk-shifting option. The big difference in value between the suboptimal base models and the optimal base model tells us the residual agency cost of risk-shifting due to the uncertainty of the risk-shifting behavior.

Figure 1.2 and Figure 1.3 displays the surfaces of optimal risk-shifting threshold XR, the relative risk-shifting incentive K, and the optimal coupon C, generated by the set of proportional financial constraints costs $i \in [0, 0.6]$ and the debt rollover rates $m \in [0, 0.2]$ which corresponds to average debt maturities of ∞ down to 5 years.

The following can be observed from these figures:

1. The optimal risk-shifting incentives decrease as financial constraints cost i increases. That is, more constrained firms are less likely to be exposed to risk-shifting problems.

2. As indicated by K, financially unconstrained firms tend to shift risk even when they are still healthy whereas financially constrained firms are more likely to shift risk when they are distressed ($K \downarrow 1$). This result provides a new perspective to look at the risk-shifting problem. As an extension of the current view that risk-shifting behaviors are more likely to happen in financially distressed firms, this model predicts that financially healthy and unconstrained firms are likely to do so as well.

3. The optimal coupon C increases as financial constraints cost i goes up for the long-term debts. It slightly decreases for the short-term debt. Moreover, C is the lowest for maturities lower than 10 years. This low leverage induces the low risk-shifting threshold XR and even lower default threshold XDL for long maturities. This explains the discrepancy we observe in the first result.

Table 1.1 presents explicitly the results generating Figure 1, 2 and 3 at representative financial constraints cost i. Panel A shows the results of the benchmark model in which no risk-shifting prospect exists. Without the agency cost of risk-shifting, the benchmark firm has an initial firm value of 1738.22, uses a coupon of 17.62 and a leverage of 59.8%, and defaults at $XD = 34.55$.

Panel B, on the other hand, shows scenarios of the firm with risk-shifting options under different proportional financial constraints costs i. When i is as low as 0% ($M = 30$)(unconstrained), the firm value is 1642.13, a reduction of 5.53% relative to the benchmark. The equity value is almost doubled and the leverage is reduced by 42.72% comparing to the benchmark case. The credit spread CSP is reduced by half when the model is switched to the benchmark. Therefore the magnitude of the agency cost of risk-shifting are prominent for the unconstrained firm. As the firm becomes more constrained, the agency cost of risk-shifting AC decreases. When $i = 50\%$, it is optimal for the firm to switch to the base model. The AC drops to 4.14% at this point and keeps decreasing as i goes up. These results complement the subsection 1.2.2 and show that financial constraints help to reduce not only the risk-shifting incentives but also the agency costs.

Panel C and D presents the results for the medium ($m = 0.1$) and short-term ($m = 0.2$) debt case respectively. The agency cost AC is much lower, the leverage ratio LR and the risk-shifting threshold XR are much higher than the long-term case. However, the patterns are the same.

1.3 Empirical Tests

The theoretical model in the previous section generates four testable hypotheses:

Hypothesis 1: Risk-shifting incentive is negatively related to financial constraints. That is, financially constrained firms are less likely to engage in risk-shifting behaviors than financially unconstrained firms.

Hypothesis 2: Financially unconstrained firms tend to risk-shift even when it is healthy. So financial distress is not the sufficient condition for risk-shifting. The power of financial distress in explaining risk-shifting will be weakened when financial constraints are controlled for.

Hypothesis 3: There is no monotonic relationship between debt maturity and risk-shifting incentive when financial constraints are not controlled for. However, short-term debt maturity helps to strengthen the effect of financial constraints in reducing risk-shifting incentives when financial constraints are controlled for.

Hypothesis 4: Since risk-shifting hurts firm value, financially unconstrained firms have lower asset returns and lower leverage than constrained firms.

To to the empirical test, I implement the regression framework in Eisdorfer (2008). In this framework, higher risk as represented by higher expected volatility implies two opposing effects on investments: the risk-shifting effect leading to higher current investment because equity holders gain from ripping off debt holders; the real options effect leading to lower current investment because the option value of waiting-to-invest is higher. If there is no risk-shifting, the investment-volatility relation should be significantly negative. However when risk-shifting takes into effect, this negative relation is weakened or even reversed. Firms with the strongest risk-shifting incentives will have a positively significant investment-volatility relation. This logic implies the following: (1) The investment-volatility relation is significantly negative for financially constrained firms, but insignificant or even significantly positive for financially unconstrained firms; (2) The investment-volatility relation is not significant for the interaction of financial distress and volatility when financial constraints are controlled for.

1.3.1 Data

I obtain the prices and returns from CRSP and the firm-level financial data from COMPUSTAT's North America Fundamental Annual tape. The 45-year sample period spans from 1965 to 2009. In order for a firm-year observation to be included, it must contain all information needed to calculate the variables in the tests. All financials are inflation adjusted to year 2004 dollar value. Firms in the financial (SIC code 5999-7000), utilities (SIC code 4899-4950) and not-for-profit (SIC code 8000 and above) industries are excluded from the tests because they are regulated and behave differently from the other firms. Leverage ratio is winsorized at 0.99 and Tobin's Q is capped at 10 to remove the influence of outliers. The final sample contains $55,044$ firm-year observations for $7,680$ firms. In addition, I obtain the NBER recession dates and the treasury rates from Federal Reserve Bank of St. Louis's FRED database.

1.3.2 Empirical Model

To gauge the differential effect of financial constraints on risk shifting behaviors empirically, I run OLS regressions of investment on expected future volatility with control of variables that are shown to be important to investments in the literature. Such baseline model can be formulated as follows:

$$\begin{aligned} I_{i,t+1} &= \alpha_i + \beta_{i,1} ExpVol_{i,t+1} + \beta_{i,2} Size_{i,t} + \beta_{i,3} Q_{i,t} + \beta_{i,4} Leverage_{i,t} + \\ & \quad \beta_{i,5} CashFlow_{i,t} + \beta_{i,6} DumRecession_t + \beta_{i,7} DefSpr_i + \beta_{i,8} r_{ft} + \epsilon_{i,t} \end{aligned} \quad (1.31)$$

The subscript i stands for the ith firm in the sample, and $t \geq 0$ stands for time. Define investment intensity $I_{i,t+1}$ as the ratio of capital expenditure to beginning-of-period total asset; $Size_{i,t}$ as the log of market value of asset (book value of debt + market value of equity), $Q_{i,t}$ as market value of asset to book value of asset, $Leverage_{i,t}$ as book value of debt over market value of assets, and $CashFlow_{i,t}$ as operating cash flow to beginning-of-period total asset. The dummy variable for NBER recession dates $DumRecession_t$ equals to 1 if the economy in any month of year t is in recession. The default spread $DefSpr_t$ is the difference between the Baa-rated and the Aaa-rated bond yields. The risk-free rate r_{ft} is taken as the 1-month Treasury bill rate.

One of the key variables in the model is the expected volatility $ExpVol_{i,t+1}$. I use the expected industry volatility as a proxy to minimize the endogenous problem in model estimation caused by the potential dependency of firm-level volatility on the other control variables. The calculation of this variable is a two-step procedure. First, the monthly value-weighted industry returns are computed based on the 2-digit SIC code for the period of 1950 to 2009. Second, a **GARCH (1,1)** model is applied to the return series of each industry separately, and the 12-month-ahead volatility forecast can be obtained for each industry at the end of each year using the model estimates.

The other key variable is the dependent variable $I_{i,t+1}$: investment intensity. I use the industry-adjusted firm-level investment intensity in the tests. In every year, investment intensities are adjusted by taking out the median investment intensity of the industry defined by the 4-digit SIC code. If in any given year, there are less than five observations in the industry, I use the 3-digit SIC code in stead to define this industry. If the same situation occurs again, the 2-digit SIC code is applied. This procedure is adopted from Eisdorfer(2008) too.

1.3.3 Financial Constraints and Financial Distress Criteria

The tests in this paper require classifying financially constrained and unconstrained firms based on some criteria. Since there is no consensus on the best criteria to use, I apply two characteristic-based and two model-based approaches that are widely used in the literature.

The first two criteria are defined with respect to rankings of firms based on size and payout ratio (i.e. dividends over net income). At the beginning of each year within the sample period, decile portfolios are constructed according to the rankings and the bottom(top) three deciles are defined as financially constrained(unconstrained) firms. The *Size* scheme follows from the arguments in Gilchrist and Himmelbert(1995) that small firms are young and less well know, thus are more vulnerable to capital market imperfections. The *Payout* scheme, on the other hand, follows from Fazarri et al. (1988) in that financially constrained firms have significantly lower payout ratios.

The next two criteria are model-based. They are good candidates because they are continuous in financial constraints and make the model predictions directly testable. The first one is the SA index introduced by Hadlock and Pierce (2010). This index is built from an ordered logit regression. It is calculated as $-(0.737*Size)+(0.043*Size^2)-(0.040*Age)$, where Size equals the log of book assets, and Age is the number of years the firm is listed with a non-missing stock price on COMPUSTAT. Size is winsorized at (the log of) 4.5 billion, and Age is winsorized at 37 years. The fact that it is based on the two relatively more exogenous variables makes it an appealing index to use. The second continuous measure is the WW index introduced by Whited and Wu (2006). This index is a linear combination of six variables computed from Compustat data. In addition to size, it also takes cash flow, dividend payout dummy, long-term leverage, industry sales growth and firm sales growth into account: WW=-0.091*CF-0.062*Div+0.021*Leverage-0.044*Size+0.102*IndustrySG-0.035*SG. All values are inflation-adjusted to 2004 dollar value. These two indices are higher if the firms are more financially constrained.

In order to define financial distress, I also use Altman's Z-score to classify the firms. Z-score = 1.2(Working capital/Total assets)+ 1.4(Retained earnings/Total assets)+3.3(Earnings before interest and taxes/Total assets)+ 0.6(Market value of equity/Book value of total liabilities) + 0.999(Sales/Total assets). A firm with Z-score lower than 1.81 at the beginning of the year is classified as in financial distress.

1.3.4 Main Results

Table 1.2 summarizes the firm characteristics for sub-samples partitioned based on the above mentioned financial constraints criteria. As seen in the table, financially constrained firms are smaller in size and have lower payout. Except for the *Payout* scheme, they have lower investments and less cash flows, and use lower leverage and shorter debt maturities on average. However, they have more investment opportunities as represented by the higher Tobin's Q, higher cash holdings, higher Z-Score and SA index. The *Payout* scheme is very different from the other three in that it generates more unbalanced classifications: only 9,312 firm-year observations are financially unconstrained, whereas the constrained group is almost four times that of the unconstrained. As a result, the pattern of the variables except size and payout are different from the other three groups.

Before testing the hypotheses, regressions of the industry-adjusted investment intensity on expected volatility for financially healthy and distressed firms are done to re-examine the results in Eisdorfer (2008) with the new sample period and the slightly differently defined control variables. Eisdorfer (2008) normalizes investment and cash flows by beginning-of-the-year PP&Es, whereas in this paper these two quantities are scaled by lagged total assets. This paper also replaces market-to-book ratio with Tobin's Q and book leverage with quasi-market leverage. The results in Table 1.4 shows that Eisdorfer (2008) 's result still holds in the new sample and with the new definitions of control variables: the financially healthy firms have a significantly negative (t-stat=-2.52) investment-volatility relationship and the financially distressed firms have a significantly positive (t-stat=2.41) investment-volatility relationship. This confirms the intuition in this regression framework: the risk-shifting effect dominates the real options effect for the financially distressed firm thus we can observe a positive regression coefficient.

Hypothesis 1

Table 1.4 examines the effect of financial constraints on reducing risk-shifting incentives. In this table, the industry-adjusted investment I_{adj} is regressed on expected industry volatility σ_{Ind} and an interactive term of σ_{Ind} and financial constraints index with the control of beginning-of-period firm size, investment opportunities (Tobin's Q), market leverage, operating cash flow, and macro level effects such as the NBER recession indicator, the default spread and the risk-free interest rate. Newey and West (1987) procedure is used to correct for heteroskedasticity and serial correlation. Since there are four financial constraints criteria considered, the financial constraints index in the interactive term takes the form of a dummy variable for the *Size* and *Payout* criteria and equals the SA and WW indices for the other two. Instead of splitting the sample into financially constrained and unconstrained sub-samples based on these criteria, I use the interactive term to introduce the financial constraints effect. This approach is intuitively appealing especially for the SA index and WW cases because it allows for a comparison of constrained and unconstrained cases on a continuous basis.

As in Eisdorfer (2008), when the real options effect dominates, the investment-volatility relation is significantly negative, whereas as the risk-shifting effect gets stronger the relation can be reversed. In this study financial constraints work as an added negative effect. This means that when the degree of financial constraint is high, the negative investment-volatility relation in Eisdorfer's (2008) framework strengthens, and the positive investment-volatility relation may be overturned. Hence the combined result of the real options effect and the financial constraints costs leads to further reduction of risk-shifting incentives. This intuition is supported by the results in Table 1.4. There exists a significant negative effect of financial constraints across all four schemes shown by the negative coefficients on the interactive terms (t-statistics of -2.77, -2.71, -2.66 and -4.02 respectively). The positive coefficients on expected volatility for the *Size* and *Payout* schemes indicate that risk-shifting effect dominates the other two effects for financially unconstrained firms. Those for the SA index and WW index schemes are negatively significant because SA and WW are negative in this sample and these coefficients indicate the negative investment-volatility relationship for financially constrained firms. Consistent with Hypothesis 1, these results show that the financially constrained firms avoid risk-shifting behaviors, whereas the unconstrained firms show some systematic tendency to engage in such activities.

Hypothesis 2

The current view of the risk-shifting literature is that risk-shifting tends to happen when firms are in financial distress. However, the theoretical model in this paper predicts that unconstrained firms are more likely to shift risk when they are still healthy, so financial distress is not as informative as before when financial constraints are controlled for. To incorporate this view into the empirical tests, I add Altman's Z-score and an interactive term of Z-score and σ_{Ind} into the regression model in Table 1.5. The results are shown in Table 1.6. In this table, the coefficients on the interactive term for financial constraints are still negatively significant once Z-score is controlled for under all four financial constraints schemes. Moreover, the interactive terms for financial distress are not significant, supporting the model prediction.

Hypothesis 3

The summary statistics in Table 1.2 show that financially constrained firms tend to have shorter debt maturities. Therefore concerns may arise based on the argument in Barnea, Haugen, and Senbet (1980) that short-term debt will be used in firms with high risk-shifting incentives to reduce such incentives because short-term debt values are less sensitive to fluctuations in volatility. This claim is supported by Barclay and Smith (1995a) and Guedes and Opler (1996) who find that growth firms use more short-term debt. However Figure 1.1 shows that there is no linear linkage between short-term debt and reduction of risk-shifting incentive unconditionally. Conditioned on financial constraints, short-term debt may help to strengthen the effect of financial constraints in reducing risk-shifting incentives. Empirically, this model prediction can be tested by introducing interactive terms between σ_{Ind} and dummy variables for different debt maturities while controlling for financial constraints. The average debt maturity is defined as the weighted average of debt maturities where the weights are proportion of debts maturing in each year[4], i.e.

[4] This definition uses the debt maturity variables in COMPUSTAT, i.e. dd1, dd2, dd3, dd4 and dd5. For debt maturing over 5 years, the maturity is defined as 6 years.

$T = \frac{1*dlc+2*dd2+3*dd3+4*dd4+5*dd5+6*(dltt-dd2-dd3-dd4-dd5)}{dlc+dltt}$. The dummy variables $D_1 = 1$ if $T <= 1$ and 0 otherwise; $D_2 = 1$ if $1 < T \leq 2$ and 0 otherwise; and D_3, D_4, D_5 can be defined in the same fashion. These dummy variables indicate whether the average debt maturity T falls into each range. The interactive terms of D_1 to D_5 with σ_{Ind} and financial constraints index FC show how debt maturities affect the effect of financial constraints on risk-shifting. If the short-term debt effect is very strong, we would observe a significantly negative coefficient on $\sigma_{Ind} * FC * D_1$ and possibly a significantly positive coefficient on $\sigma_{Ind} * FC * D_5$. Table 1.6 presents results from such tests.

The first column in Table 1.6 shows the effect of debt maturity on risk-shifting incentives without control of financial constraints. The interactive terms for this column are actually between σ_{Ind} and $D_i, i = 1, 2, ...5$ only. The insignificant coefficients on the interactive terms shows no sign of risk-shifting reduction effect. The next four columns present the same results with control of financial constraints. The 1-year debt significantly helps to strengthen the effect of financial constraints as seen by the negative coefficients of $\sigma_{Ind} * FC * D_1$ for all criteria except *Payout*(t-stat=-5.02, -4.10, -4.50). The 2-year debt maturity helps as well, but the effect is less significant. Surprisingly, the 3-year and 4-year debt contribute more to increases in risk-shifting incentive than the 5-year case, but this result is not supported by the *WW* index. This result is consistent with the view that short-term debt maturity alleviates risk-shifting behaviors, but this works only conditionally when financial constraints are controlled for. As a result, debt maturity is more of a complement than substitute to financial constraints for reducing risk-shifting incentives.

Hypothesis 4

The tests in the previous subsections show that financial constraints have a significant effect on reducing risk-shifting incentives. A natural question following this result is how big the impact is. Table 1.7 addresses this question. As discussed in section 1.2.2, the agency cost of risk-shifting can be measured by the reduction of the optimal leverage and the loss of firm value relative to the benchmark. Financial constraints help in such a way that the reduction is lessened as firms become more constrained. To illustrate this pattern, I run regressions of industry-adjusted leverage and asset returns on contemporaneous industry-adjusted investment I_{adj}, interactions of I_{adj} with financial constraints variable FC and other control variables in the baseline model. The first dependent variable, the industry-adjusted leverage, is computed as the firm-level leverage minus the median industry leverage in the corresponding year based on the 4-digit SIC code to control the industry effect (see Frank and Goyal (2009)). The second dependent variable, the industry-adjusted asset return, is computed similarly where asset return is defined as the percentage change of the market value of asset(book value of debt plus market value of equity). According to *Hypothesis 4*, the more constrained firms should have less reductions in leverage and firm value, so the coefficient of the interactive terms should be significantly positive.

The first four columns in Table 1.7 present the regression results for the industry-adjusted leverage under the four financial constraints criteria. Except for the *Payout* scheme, financial constraints help to reserve the level of leverage under all schemes (t-statistics are 4.58, 3.64 and 5.76 respectively for the $FC * I_{adj}$ term). The coefficient is negatively significant under the *Payout* scheme because the classification under this scheme is highly skewed and the constrained sub-sample is more likely to be contaminated with unconstrained firms. According to the *Size* scheme, everything else being equal, financially constrained firms should be able to use 0.152% more leverage for every 1% investment they make compared to unconstrained firms.

The next four columns in Table 1.7 shows the regression results for the industry-adjusted asset return. Everything else being equal, the asset returns of the financially constrained firms are 0.605% higher than that of the unconstrained firms on average for every 1% investment they make based on *Size* scheme. This implies that financially constrained firms make better investment choices by avoiding risk-shifting behaviors. The more investment they make, the greater the difference they can make. Similar conclusion can be drawn under the other three schemes.

1.4 Robustness Check

This study assumes that a firm requires external financing when equity holders implement risk-shifting strategies. What if the firm has enough internal funding so that they can shift risk with a much lower cost? To address this concern, let's first look at the summary statistics in Table 1.2. Cash holdings *CH* as defined by total cash normalized by lagged total asset represents the amount of internal funding for a firm. On average, unconstrained firms hold much less cash, so they will resort to external financing when they make a major investment. Even if the internal fundings are enough, the low cost of external financing makes it a legitimate thing to do. On the other hand, financially constrained firms have more cash holdings. However, the difficulty of acquiring external financing is the exact reason that they preserve more internal fundings. It is never a good idea for equity holders of the constrained firms to exhaust the internal fundings and make their risk-shifting attempt so obvious. Therefore, it is reasonable to assume that risk-shifting behaviors are subject to financial constraints costs potentially. Table 1.8 presents the results of the baseline model in Table 1.5 with added control of *CH*. The significantly negative coefficients on $\sigma_{Ind} * FC$ (t-stat=-2.82, -2.76, -2.56 and -3.87) show that financial constraints effect still persists even when *CH* in controlled for. Overall, even when firm has abundant internal fundings, financial constraints will still lead to less risk-shifting incentives.

1.5 Conclusion

Risk-shifting has been a center of debate in corporate finance. Various studies have provided evidences on the existence of risk-shifting, the quantitative effect of it and on how to reduce it. This study addresses all three problems from a novel perspective by looking at the link between financing frictions and corporate investment via the risk channel. When financial constraints are linked to risk-shifting investment, it can help reduce risk-shifting incentives. That is, firms are more likely to shift risk when they are financially unconstrained. Both the theoretical implications and empirical tests support this result. Moreover, this study refreshes the current view that firms shift risk when they are financially distressed. In fact, even the healthy firms engage in risk-shifting behaviors when they are financially unconstrained. This differentiation of potential risk-shifting and non risk-shifting targets can generate interesting implications for the simultaneous choice of capital structure and investment decisions as well as asset pricing, which will be explored in future research. This study also find that short-term debt does not reduce risk-shifting incentives unconditionally; it helps to strengthen the effect of financial constraints. Another interesting finding is that the quantitative effect of financial constraints in reducing risk-shifting behavior is very high. A constrained firm typically generates 0.605% more return and uses 0.152% more leverage compared with a similar unconstrained firm for every 1% investment they make. The effect can be big when firms make massive investments. Along this line, it would be interesting to study how this effect breaks down to equity returns and bond returns. I will leave it as a future research topic.

1.6 References

Almeida, Heitor, Murillo Campello, and Michael Weisbach, 2004, The Cash Flow Sensitivity of Cash, *Journal of Finance 59*, 1777-804.

Almeida, Heitor, and Murillo Campello, 2007, Financial Constraints, Asset Tangibility, and Corporate Investment, *Review of Financial Studies 20(5)*, 1429-1460.

Altman, Edward I., 1968, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance 23*, 589-609.

Altman, Edward I., 2000, Predicting Financial Distress of Companies: Revisiting the Z-score and Zeta Models, Working paper, New York University.

Barclay, Michael J., and Clifford W. Smith, 1995a, The Maturity Structure of Corporate Debt, *Journal of Finance 50*, 609-631.

Barnea, Amir, Robert A. Haugen, and Lemma W. Senbet, 1980, A Rationale for Debt maturity Structure and Call Provisions in an Agency Theoretic Framework, *Journal of Finance 35*, 1223-1234.

Campello, Murillo, and Dirk Hackbarth, 2008, Corporate Financing and Investment: the Firm-level Credit-multiplier, Working paper, University of Illinois, Urbana-Champaign.

Campello, Murillo, Graham, John R. and Harvey, Campbell R., 2010, The Real Effects of Financial Constraints: Evidence from a Financial Crisis, *Journal of Financial Economics 97(3)*, 470-487.

Dixit, A.,1993,The Art of Smooth Pasting, *Fundamentals of Pure and Applied Economics.*

Dixit, A., R. Pindyck, 1994, Investment under uncertainty, Princeton: Princeton University Press.

Eisdorfer, Assaf, 2008, Empirical Evidence of Risk Shifting in Financially Distressed Firms, *Journal of Finance 63 (2)*, 609-637.

Ericsson, Jan, 2000, Asset Substitution, Debt Pricing, Optimal Leverage and Maturity, *Finance 21*, 39-70.

Fazzari, Steven M., Glenn R. Hubbard, and Bruce C. Petersen, 1988, Finance Constraints and Corporate Investment, *Brookings Papers on Economic Activity 1*, 141-195.

Frank, Murray Z., Vidhan K. Goyal, 2009, Capital Structure Decisions: Which Factors Are Reliably Important? *Financial Management (Blackwell Publishing Limited)38(1)*, 1-37

Galai, Dan, and Ronald W. Masulis, 1976, The Option Pricing Model and the Risk Factor of Stock, *Journal of Financial Economics 3*, 53-81.

Goldstein, Robert, Nengjiu Ju and Hayne Leland, 2001, An EBIT-based Model of Dynamic Capital Structure, *Journal of Business 74(4).*

Hadlock, Charles J., and Joshua R. Pierce, 2010, New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index, *The Review of Financial Studies 23*,1909-1940.

Hennessy, Christopher A., and Yuri Tserlukevich, 2004, Dynamic Hedging Incentives, Debt, and Warrants, *Working paper, University of California, Berkeley.*

Hennessy, Christopher A. and Whited, Toni M., 2007, How Costly Is External Financing? Evidence from a Structural Estimation, *Journal of Finance 62*, 1705-1745.

Jensen, Michael C., and William H. Meckling, 1976, Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure, *Journal of Financial Economics 3*, 305-360.

Leland, Hayne E., 1998, Agency Costs, Risk Management, and Capital Structure, *Journal of Finance 53*, 1213-1243.

McDonald, Robert, and Daniel Siegel, 1986, The Value of Waiting to Invest, *Quarterly Journal of Economics 101*, 707-727.

Merton, Robert C., 1973, Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science (The RAND Corporation) 4 (1)*, 141-183.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple Positive Semidefinite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica 55*, 703-708.

Smith, Clifford W., and Jerold B. Warner, 1979, On Financial Contracting: An Analysis of Bond Covenants, *Journal of Financial Economics 7*, 117-162.

Subramanian, Ajay, 2003, Carrots or Sticks? Optimal Compensation for Firm Managers, Working paper, Georgia Institute of Technology.

Whited, Toni, and Wu, Guojun, 2006, Financial Constraints Risk, *The Review of Financial Studies 19 (2)*, 531-559

1.7 Tables and Figures

Table 1.1: Optimal Risk-shifting Thresholds over a Set of Financial Constraints Costs and Debt Maturities

This table presents the effect of financial constraints cost $M = M_0(1+i)$ on reducing risk-shifting incentives and agency costs of risk-shifting for a representative firm. The fixed cost $M_0 = 30$ is the cost of a risk-shifting investment for a financially unconstrained firm and the percentage financial constraints cost is denoted by $i \in [0, 0.55]$ which increases with degree of financial constraints. This table also presents the effect for three exogenously given debt rollover rates: $m \in 0, 0.1, 0.2$, which is equivalent to average debt maturities of ∞, 10 years and 5 years. **Panel A** shows the results for the benchmark model in which risk-shifting can be credibly excluded ex ante, whereas **Panel B, C and D** shows the results for the risk-shifting model in which risk-shifting option can be exercised ex post for different debt maturities. The optimal coupon for the benchmark model is denoted as C^*. The other quantities for the benchmark model,i.e. optimal firm value VB, equity value SB, leverage ratio LR, default threshold XD and credit spread CSP are all evaluated at C^*. For the risk-shifting model, the optimal coupon C is determined at time 0, the initial firm value $V0$, the initial equity value $S0$ and the initial equity value evaluated with the benchmark model $SB(C)$ are all computed at the optimal coupon C in the risk-shifting model. The leverage ratio LR is computed as initial debt value over firm value. The ex post optimal risk-shifting threshold, default threshold after risk-shifting and default threshold before risk-shifting are denoted by XR, XDH and XDL respectively. The relative risk-shifting incentive is $K = XR/XDL$. The agency cost $AC = (V0 - VB)/VB$ is percentage loss of firm value relative to the benchmark model. The credit spread $CSP = C/(V0 - S0) - r_f$ is expressed as basis points. The firm has an initial cash flow $X_0 = 100$.The marginal tax rate $\tau = 20\%$ and the expected rate of return $\mu = 1\%$. The pre- and post-shifting risk levels are $\sigma_L = 25\%$ and $\sigma_H = 40\%$. The percentage cost of bankruptcy is $\alpha = 0.25$ and the risk-free rate is $r_f = 6\%$.

Panel A Benchmark Model, m=0 (T=Inf)

C*	VB(C*)	SB(C*)	LR(C*)	XD(C*)	CSP(C*)
79.62	1,738.22	698.71	59.80%	34.55	165.9

Panel B Risk-shifting Model, m=0 (T=Inf)

i	C	V0	S0	SB(C)	LR	XR	XDH	XDL	K	AC	CSP
0%	18.58	1,642.13	**1,361.71**	1,359.95	17.08%	35.86	5.38	8.06	4.45	5.53%	62.6
5%	19.03	1,643.14	**1,355.96**	1,354.36	17.48%	35.04	5.51	8.26	4.24	5.47%	62.6
10%	19.47	1,644.14	**1,350.35**	1,348.91	17.87%	34.28	5.64	8.45	4.06	5.41%	62.6
15%	19.90	1,645.11	**1,344.86**	1,343.60	18.25%	33.58	5.76	8.63	3.89	5.36%	62.6
20%	20.31	1,646.06	**1,339.49**	1,338.42	18.62%	32.91	5.88	8.82	3.73	5.30%	62.6
25%	20.72	1,646.98	**1,334.24**	1,333.35	18.99%	32.28	6.00	8.99	3.59	5.25%	62.6
30%	21.12	1,647.89	**1,329.10**	1,328.40	19.35%	31.69	6.12	9.17	3.46	5.20%	62.6
35%	21.52	1,648.79	**1,324.05**	1,323.54	19.70%	31.14	6.23	9.34	3.33	5.14%	62.6
40%	21.90	1,649.66	**1,319.10**	1,318.79	20.04%	30.61	6.34	9.51	3.22	5.09%	62.6
45%	22.28	1,650.52	**1,314.24**	1,314.13	20.37%	30.11	6.45	9.67	3.11	5.05%	62.6
50%	22.66	1,666.29	1,309.56	1,309.56	21.41%	29.63	6.56	9.83	3.01	4.14%	35.1
55%	23.02	1,667.19	1,305.08	1,305.08	21.72%	29.18	6.67	9.99	2.92	4.09%	35.8

Panel C Risk-shifting Model, m=1/10 (T=10)

i	C	V0	S0	SB(C)	LR	XR	XDH	XDL	K	AC	CSP
0%	32.82	1,651.33	**1,170.82**	1,162.05	29.10%	51.56	15.30	20.70	2.49	2.37%	83.1
5%	32.84	1,651.50	**1,169.58**	1,161.83	29.18%	50.37	15.33	20.71	2.43	2.36%	81.4
10%	32.86	1,651.66	**1,168.38**	1,161.59	29.26%	49.26	15.36	20.72	2.38	2.35%	79.9
15%	32.88	1,651.82	**1,167.21**	1,161.35	29.34%	48.20	15.39	20.73	2.33	2.34%	78.4
20%	32.90	1,651.98	**1,166.07**	1,161.10	29.41%	47.19	15.42	20.74	2.28	2.33%	77.0
25%	32.92	1,652.13	**1,164.96**	1,160.83	29.49%	46.24	15.45	20.75	2.23	2.32%	75.7
30%	32.94	1,652.27	**1,163.88**	1,160.57	29.56%	45.34	15.48	20.76	2.18	2.31%	74.4
35%	32.96	1,652.42	**1,162.83**	1,160.30	29.63%	44.47	15.51	20.78	2.14	2.31%	73.2
40%	32.98	1,652.55	**1,161.81**	1,160.02	29.70%	43.65	15.54	20.79	2.10	2.30%	72.1
45%	33.00	1,652.69	**1,160.81**	1,159.74	29.76%	42.86	15.57	20.80	2.06	2.29%	71.0
50%	33.03	1,652.82	**1,159.83**	1,159.46	29.83%	42.11	15.60	20.82	2.02	2.28%	69.9
55%	33.05	1,674.99	1,159.18	1,159.18	30.80%	41.39	15.62	20.83	1.99	0.97%	40.7

Panel D Risk-shifting Model, m=1/5 (T=5)

M	C	V0	S0	SB(C)	LR	XR	XDH	XDL	K	AC	CSP
0%	32.75	1,648.32	**1,147.98**	1,146.21	30.35%	44.19	18.95	24.04	1.84	1.35%	54.6
5%	32.59	1,648.22	**1,149.25**	1,148.37	30.27%	43.35	18.89	23.93	1.81	1.36%	53.2
10%	32.44	1,648.13	**1,150.48**	1,150.45	30.19%	42.55	18.84	23.83	1.79	1.36%	51.8
15%	32.29	1,665.20	1,152.43	1,152.43	30.79%	41.77	18.78	23.73	1.76	0.34%	29.6
20%	32.14	1,665.09	1,154.34	1,154.34	30.67%	41.03	18.73	23.63	1.74	0.35%	29.3
25%	32.01	1,664.98	1,156.17	1,156.17	30.56%	40.32	18.67	23.54	1.71	0.35%	29.0
30%	31.87	1,664.87	1,157.93	1,157.93	30.45%	39.63	18.62	23.45	1.69	0.36%	28.8
35%	31.75	1,664.77	1,159.63	1,159.63	30.34%	38.97	18.58	23.37	1.67	0.37%	28.5
40%	31.63	1,664.67	1,161.26	1,161.26	30.24%	38.34	18.53	23.29	1.65	0.37%	28.2
45%	31.51	1,664.57	1,162.84	1,162.84	30.14%	37.73	18.48	23.21	1.63	0.38%	28.0
50%	31.39	1,664.48	1,164.37	1,164.37	30.05%	37.13	18.44	23.13	1.61	0.38%	27.7
55%	31.28	1,664.39	1,165.84	1,165.84	29.95%	36.56	18.40	23.06	1.59	0.39%	27.5

Table 1.2: Summary Statistics

This table presents summary statistics for full and sub-samples of financially constrained and unconstrained firms under four financial constraints schemes: *Size*, *Payout*, *SA* and *WW*. Book value of asset *TA* is in millions of dollars. Industry expected volatility σ_{Ind} is the expected industry volatility for the year ahead from a $GARCH(1,1)$ model based on stock returns over the period of 1950 to 2009. Tobin's Q is defined as market value of assets over book value of assets. Leverage ratio *LR* is the ratio of book value of debt over market value of assets. Investment intensity *I*, cash flow *CF*, and Cash Holding *CH* are defined as the ratio of capital expenditures, operating cash flows and total cash to beginning-of-period total assets. Payout ratio *Payout* is defined as dividends and stock repurchases over operating income. *Z-Score* is introduced in Altman (1968) for forecasting bankruptcy. The Size-Age index *SA* is introduced in Hadlock and Pierce (2010), and the Whited-Wu index *WW* is introduced in Whited and Wu (2006). They both are continuous measure of financial constraints. Average debt maturity *T* is the weighted average of debts maturing in one, two, three, four and five years. *N* is the number of observations in each sample. All financials are inflation-adjusted to 2004 dollar value. The sample data is taken from CRSP for the period of 1950 to 2009 and COMPUSTAT for the period of 1965 to 2009. Financial, utility and not-for profit firms are excluded. Leverage ratio is winsorized at 0.99. Q is capped at 10. The top and bottom percentiles of *I*, *CF*, *AT*, *Z-Score* and *WW* are excluded to remove outliers. The final full sample contains 7,680 firms.

			Size		**Payout**		**SA**		**WW**	
		Full Sample	**Const**	**Unconst**	**Const**	**Unconst**	**Const**	**Unconst**	**Const**	**Unconst**
	Mean	889.545	36.720	2,618.717	571.253	1,049.521	42.553	306.995	160.581	237.834
TA	Median	174.724	29.949	1,409.910	123.572	176.719	33.284	91.620	35.846	87.438
	Std. Dev.	2,030.922	30.033	3,064.964	1,448.575	2,379.024	32.311	809.892	738.240	680.643
	Mean	0.217	0.215	0.214	0.218	0.218	0.216	0.219	0.219	0.219
σ_{Ind}	Median	0.201	0.198	0.200	0.202	0.200	0.199	0.202	0.200	0.202
	Std. Dev.	0.066	0.065	0.066	0.066	0.070	0.066	0.067	0.073	0.067
	Mean	1.832	2.069	1.693	1.872	1.936	2.169	1.929	2.116	1.901
Q	Median	1.426	1.553	1.387	1.427	1.515	1.660	1.478	1.597	1.454
	Std. Dev.	1.246	1.514	1.001	1.308	1.321	1.551	1.357	1.539	1.342
	Mean	0.168	0.117	0.222	0.181	0.123	0.116	0.157	0.124	0.154
LR	Median	0.116	0.051	0.185	0.123	0.065	0.048	0.089	0.055	0.087
	Std. Dev.	0.176	0.149	0.178	0.187	0.152	0.151	0.180	0.155	0.177
	Mean	0.068	0.061	0.072	0.073	0.054	0.069	0.069	0.064	0.068
I	Median	0.044	0.034	0.052	0.044	0.036	0.039	0.042	0.035	0.041
	Std. Dev.	0.075	0.077	0.068	0.082	0.063	0.083	0.080	0.081	0.080
	Mean	0.060	-0.00012	0.102	0.046	0.040	-0.005	0.042	-0.011	0.040
CF	Median	0.078	0.034	0.099	0.066	0.068	0.031	0.065	0.027	0.063
	Std. Dev.	0.158	0.207	0.093	0.168	0.177	0.216	0.178	0.209	0.175
	Mean	0.185	0.180	0.204	0.011	0.776	0.163	0.166	0.161	0.165
Payout	Median	0.026	0.000	0.110	0.000	0.971	0.000	0.000	0.000	0.000
	Std. Dev.	0.303	0.342	0.262	0.024	0.260	0.331	0.315	0.334	0.316
	Mean	0.224	0.305	0.118	0.243	0.277	0.350	0.272	0.334	0.269
CH	Median	0.088	0.156	0.048	0.092	0.143	0.175	0.118	0.167	0.120
	Std. Dev.	0.449	0.534	0.260	0.481	0.534	0.629	0.518	0.600	0.509
	Mean	4.789	5.242	3.863	4.778	4.973	5.609	5.101	5.111	5.056
Z-Score	Median	3.608	3.912	3.126	3.449	3.962	4.094	3.741	3.644	3.728
	Std. Dev.	4.457	5.217	3.179	4.725	4.555	5.498	4.924	5.352	4.862
	Mean	-3.164	-2.412	-3.903	-2.969	-3.230	-2.285	-2.761	-2.518	-2.850
SA	Median	-3.119	-2.376	-3.868	-2.950	-3.191	-2.364	-2.834	-2.479	-2.845
	Std. Dev.	0.772	0.518	0.543	0.693	0.810	0.347	0.499	0.594	0.622
	Mean	-0.147	-0.035	-0.263	-0.115	-0.156	-0.039	-0.103	0.014	-0.084
WW	Median	-0.157	-0.059	-0.279	-0.128	-0.170	-0.061	-0.119	-0.029	-0.112
	Std. Dev.	0.142	0.118	0.109	0.137	0.146	0.119	0.130	0.127	0.122
	Mean	1.360	1.132	1.473	1.405	1.123	1.152	1.299	1.158	1.299
T	Median	1.232	1.000	1.319	1.298	1.000	1.009	1.131	1.000	1.153
	Std. Dev.	1.101	1.011	1.063	1.112	1.093	1.006	1.122	1.063	1.112
N		55,044	16,524	16,524	33,281	9,312	16,514	38,530	16,514	38,530

Table 1.3: Regressions of Investment on Expected Volatility for Financially Healthy and Distressed Firms

This table presents results from the regressions of investment on expected volatility for the financially healthy and distressed firms. Firms with Altman's Z-score no more than 1.81 are classified as financially distressed; the rest of them are healthy. Column *Healthy* represent regressions for the healthy firms and *Distress* for the distressed firms. Investment intensity I is defined as the ratio of capital expenditures to beginning-of-period total assets. The dependent variable is the industry-adjusted investment intensity I_{adj} at the beginning of the year, which is the firm-level investment intensity minus the median investment intensity of the industry in the corresponding year.σ_{Ind} is the expected industry volatility at the beginning of the year estimated by applying GARCH (1,1) model to the value-weighted industry returns defined by the 2-digit SIC codes. *Size* is the log of total market value of assets which equals to the market value of equity and the book value of debt. Tobin's Q is the market value of assets over the book value of assets. Leverage LR is the book value of debt over market value of assets. The cash flow CF is the ratio of operating cash flow to beginning-of-period total assets. *Recession* is the dummy variable which equals to one if the firm-year is in recession according to NBER's recession dates. *DefSpr* is the spread between Baa-rated and Aaa-rated long term bond yields. Risk-free Rate r_f is the 1-month T-bill rate. N represents the number of observations used in the regressions. The sample data is taken from CRSP and COMPUSTAT for the period of 1965 to 2009. Financials, utilities and not-for-profit firms are excluded. The full sample contains 7,680 firms. The regression coefficients and t-statistics are calculated based on Newey-west standard errors with 6 lags.

	Healthy	**Distressed**
σ_{Ind}	**-0.014**	**0.026**
	(-2.52)*	**(2.41)***
Size	-0.001	0.000
	(-2.70)**	(-0.15)
Q	0.009	0.011
	(22.30)**	(10.58)**
LR	-0.01	-0.012
	(-3.28)**	(-2.74)**
CF	0.061	0.04
	(20.16)**	(6.96)**
Recession	-0.003	-0.001
	(-4.06)**	(-0.49)
DefSpr	0.186	0.398
	(2.25)*	(2.39)*
r_f	0.132	-0.022
	(7.44)**	(-0.57)
Intercept	-0.005	-0.021
	(-2.00)*	(-4.26)**
N	45,382	9,663

* $p < 0.05$; ** $p < 0.01$

Table 1.4: Regressions of Investment on Expected Volatility and its Interaction with Financial Constraints Index

This table presents results from the regressions of investment on expected volatility and its interaction with financial constraints index. The four columns *Size*, *Payout*, *SA* and *WW* shows the four financial constraints classification schemes. The first two are based on the ranking of book size (the log of book value of assets) and payout ratio (dividends and stock repurchases over operating income). Firms that are in the bottom (top) three deciles are defined as financially constrained (unconstrained). The next two are continuous in financial constraints. *SA* index is introduced in Hadlock and Pierce (2010). It includes the linear terms of size and firm age as well as a quadratic term of size in the model. *WW* index is introduced in Whited and Wu (2006) and is a linear combination of cash flow, dividend payout dummy, total long-term debt, total asset and sales growth. The dependent variable is the industry-adjusted investment intensity I_{adj} at the beginning of the year. The independent variables are the expected industry volatility σ_{Ind} at the beginning of the year, the financial constraints index FC and the interaction of FC and σ_{Ind}. FC is a dummy variable which equals to 1 if the firm is financially constrained and 0 otherwise for the *Size* and *Payout* scheme whereas it equals *SA* index or *WW* index for the other two schemes. The other control variables are defined in *Table 1.3.* The sample data is taken from CRSP and COMPUSTAT for the period of 1965 to 2009. Only non-financial and non-utility firms are included. The full sample contains 7,680 firms. The regression coefficients and t-statistics are calculated based on Newey-west standard errors with 6 lags.

	Size	**Payout**	**SA**	**WW**
σ_{Ind}	0.0110	0.0150	-0.0680	-0.0190
	(1.29)	(1.58)	(-3.00)**	(-2.97)**
FC	0.0090	0.0190	0.0140	0.0120
	(2.50)*	(7.58)**	(8.27)**	(1.7)
σ_{Ind} * **FC**	**-0.0350**	**-0.0290**	**-0.0180**	**-0.1130**
	(-2.77)**	**(-2.71)****	**(-2.66)****	**(-4.02)****
Size	0.0000	0.0010	0.0030	-0.0010
	(0.04)	(2.22)*	(8.13)**	(-3.90)**
M2B	0.008	0.009	0.007	0.009
	(16.04)**	(21.59)**	(17.56)**	(23.83)**
LR	-0.0210	-0.0250	-0.0180	-0.0180
	(-6.83)**	(-9.31)**	(-7.42)**	(-7.49)**
CF	0.0520	0.0570	0.0620	0.0580
	(16.14)**	(20.01)**	(22.43)**	(20.63)**
NBER	-0.0020	-0.0030	-0.0020	-0.0030
	(-2.34)*	(-3.36)**	(-3.19)**	(-4.09)**
DefSpr	0.3530	0.3240	0.2270	0.1940
	(3.55)**	(3.50)**	(3.00)**	(2.54)*
r_f	0.133	0.139	0.105	0.114
	(6.27)**	(7.30)**	(6.46)**	(6.98)**
Intercept	-0.014	-0.026	0.023	-0.002
	(-3.01)**	(-8.55)**	(4.17)**	(-1.07)
N	33,050	42,593	55,045	55,045

* $p < 0.05$; ** $p < 0.01$

Table 1.5: Regressions of Investment on Expected Volatility and its Interaction with Financial Constraints Index and Z-Score

This table presents results from the regressions of investment on expected volatility and its interaction with financial constraints index and Altman's Z-Score. The four columns *Size*, *Payout*, *SA* and *WW* shows the four financial constraints classification schemes. The details about the four schemes are described in *Table 1.5*. The dependent variable is the industry-adjusted investment intensity I_{adj} at the beginning of the year. The independent variables are the expected industry volatility σ_{Ind} at the beginning of the year, the financial constraints index *FC*, the Altman's *Z-Score* and the interaction of σ_{Ind} with *FC* and *Z-Score*. *FC* is a dummy variable which equals to 1 if the firm is financially constrained and 0 otherwise for the *Size* and *Payout* scheme whereas it equals *SA* index or *WW* index for the other two schemes. The Altman's *Z-Score* is based on Altman's (1968) model. The other control variables are defined in *Table 1.3*. The sample data is taken from CRSP and COMPUSTAT for the period of 1965 to 2009. Only non-financial and non-utility firms are included. The full sample contains 7,680 firms. The regression coefficients and t-statistics are calculated based on Newey-west standard errors with 6 lags.

	Size	**Payout**	**SA**	**WW**
σ_{Ind}	0.015	0.021	-0.059	-0.015
	(1.51)	(1.92)	(-2.50)*	(-1.77)
FC	0.008	0.019	0.014	0.01
	(2.24)*	(7.52)**	(8.05)**	(1.46)
σ_{Ind}***FC**	**-0.034**	**-0.029**	**-0.016**	**-0.11**
	(-2.71)**	**(-2.66)****	**(-2.41)***	**(-3.84)****
Z-Score	0.00002	0.00003	-0.00005	-0.000138
	(0.06)	(0.11)	(-0.21)	(-0.54)
σ_{Ind}***Z-Score**	-0.002	-0.002	-0.001	-0.001
	(-1.11)	(-1.44)	(-1.05)	(-0.88)
SIZE	-0.00016	0.001	0.003	-0.001
	(-0.3)	(2.40)*	(8.23)**	(-3.91)**
M2B	0.008	0.009	0.007	0.01
	(14.75)**	(19.95)**	(16.37)**	(22.07)**
LR	-0.027	-0.031	-0.024	-0.025
	(-8.76)**	(-11.57)**	(-9.88)**	(-10.22)**
CF	0.053	0.057	0.062	0.058
	(15.58)**	(19.31)**	(21.72)**	(19.99)**
NBER	-0.002	-0.002	-0.002	-0.002
	(-2.24)*	(-2.86)**	(-2.82)**	(-3.72)**
DefSpr	0.334	0.307	0.222	0.187
	(3.40)**	(3.34)**	(2.92)**	(2.45)*
r_f	0.137	0.144	0.109	0.118
	(6.38)**	(7.47)**	(6.66)**	(7.17)**
Intercept	-0.012	-0.027	0.022	-0.002
	(-2.50)*	(-7.80)**	(3.88)**	(-0.68)
N	32,184	41,670	53,783	53,783

* $p < 0.05$; ** $p < 0.01$

Table 1.6: Regressions of Investment on Expected Volatility and its Interactions with Financial Constraints Index and Debt Maturity Dummies

This table presents results from the regressions of investment on expected volatility and its interaction with financial constraints index and debt maturities. The four columns *Size*, *Payout*, *SA* and *WW* shows the four financial constraints classification schemes which are described in detail in *Table 1.4*. The first column does not take financial constraints into account. The dependent variable is the industry-adjusted investment intensity I_{adj} at the beginning of the year. The independent variables are the expected industry volatility σ_{Ind} at the beginning of the year, the financial constraints index FC, the debt maturity dummies D_1 to D_5 and the interaction of σ_{Ind} with FC and D_i, $i=1,2,3,4,5$. Please note that for the first column, these interactions are between σ_{Ind} and D_i, $i=1,2,3,4,5$ only. FC equals to 1 if the firm is financially constrained and 0 otherwise for the *Size* and *Payout* scheme whereas it equals SA index or WW index for the other two schemes. The debt maturity dummy variables D_1 up to D_5 are defined based on the weighted average debt maturity $T=(dlc*1+dd2*2+dd3*3+dd4*4+dd5*5+(dltt\text{-}dd2\text{-}dd3\text{-}dd4\text{-}dd5)*6)/(dlc+dltt)$. $D_1=1$ if $T \leq 1$ and $D_2=1$ if $1 < T \leq 2$, and so on. The variables related to debt maturities and their interactions with σ_{Ind} are omitted. The rest of the control variables are defined in *Table 1.3*. The sample data is taken from CRSP and COMPUSTAT for the period of 1965 to 2009. Only non-financial and non-utility firms are included. The full sample contains 7,680 firms. The regression coefficients and t-statistics are calculated based on Newey-west standard errors with 6 lags.

		Size	**Payout**	**SA**	**WW**
σ_{Ind}	-0.013	-0.009	0.014	-0.039	-0.016
	(-1.43)	(-0.65)	(-1.11)	(-1.56)	(-1.67)
FC		0.006	0.018	0.013	0.008
		(1.66)	(7.10)**	(7.72)**	(1.19)
σ_{Ind}***FC**		-0.014	-0.036	-0.008	-0.053
		(-0.97)	(-3.03)**	(-1.06)	(-1.57)
σ_{Ind}***FC***D_1	0.003	-0.048	0.008	-0.02	-0.117
	(0.31)	(-5.02)**	(0.99)	(-4.10)**	(-4.50)**
σ_{Ind}***FC***D_2	0.009	-0.019	0.009	-0.011	-0.055
	(0.78)	(-1.94)	(1.05)	(-2.15)*	(-1.93)
σ_{Ind}***FC***D_3	0.023	0.031	0.029	0.022	0.034
	(1.49)	(2.28)*	(2.40)*	(3.08)**	(0.9)
σ_{Ind}***FC***D_4	0.032	0.06	0.001	0.037	0.06
	(1.44)	(2.25)*	(0.04)	(3.13)**	(0.89)
σ_{Ind}***FC***D_5	-0.03	0.041	-0.007	0.006	0.049
	(-0.99)	(1.2)	(-0.25)	(0.41)	(0.55)
Size	-0.0005	-0.0003	0.001	0.003	-0.001
	(-2.15)*	(-0.54)	(2.21)*	(8.18)**	(-3.88)**
M2B	0.009	0.008	0.009	0.007	0.009
	(24.15)**	(15.93)**	(21.72)**	(17.60)**	(23.96)**
LR	-0.022	-0.028	-0.03	-0.024	-0.023
	(-9.15)**	(-9.01)**	(-10.91)**	(-9.67)**	(-9.45)**
CF	0.059	0.052	0.057	0.061	0.057
	(21.72)**	(16.11)**	(20.11)**	(22.38)**	(20.62)**
NBER	-0.002	-0.002	-0.003	-0.002	-0.002
	(-3.79)**	(-2.25)*	(-3.34)**	(-3.15)**	(-4.03)**
DefSpr	0.256	0.376	0.325	0.245	0.21
	(3.39)**	(3.78)**	(3.49)**	(3.23)**	(2.74)**
r_f	0.109	0.123	0.134	0.097	0.109
	(6.71)**	(5.82)**	(7.05)**	(5.95)**	(6.66)**
Intercept	-0.002	-0.005	-0.021	0.023	0.002
	(-0.84)	(-0.92)	(-5.82)**	(4.13)**	(0.54)
N	55,044	33,048	42,593	55,044	55,044

* $p < 0.05$; ** $p < 0.01$

Table 1.7: Regressions of Industry-adjusted leverage and Asset Return on Contemporaneous Investment and Financial Constraints

This table shows the results from the regressions of industry-adjusted leverage and asset return on contemporaneous industry-adjusted investment as well as its interaction with financial constraints index. The four columns *Size*, *Payout*, *SA* and *WW* shows the four financial constraints classification schemes which are described in detail in *Table 1.5*. Investment intensity I is defined as the ratio of capital expenditures to beginning-of-period total assets. The dependent variable, industry-adjusted leverage, is computed as firm-level leverage minus the median leverage of the industry in the same year defined by 4-digit SIC codes. Asset return, on the other hand, is the percentage change of the market value of asset (market value of equity + book value of debt). Investment intensity I is defined as the ratio of capital expenditures to beginning-of-period total assets. The industry-adjusted investment intensity I_{adj}, which is the firm-level investment intensity minus the median investment intensity of the industry in the corresponding year. The financial constraints index FC is a dummy variable which equals to 1 if the firm is financially constrained and 0 otherwise for the *Size* and *Payout* scheme whereas it equals *SA* index or *WW* index for the other two schemes. The other control variables are defined in *Table 1.3*. The sample data is taken from CRSP and COMPUSTAT for the period of 1965 to 2009. Only non-financial and non-utility firms are included. The full sample contains 7,680 firms. The regression coefficients and t-statistics are calculated based on Newey-west standard errors with 6 lags.

	Industry-adjusted Leverage				**Industry-adjusted Asset Return**			
	Size	**Payout**	**SA**	**WW**	**Size**	**Payout**	**SA**	**WW**
Iadj	-0.101	0.006	0.203	0.032	1.423	1.504	3.703	1.881
	(-4.66)**	(0.22)	(2.87)**	(-1.52)	(16.10)**	(9.71)**	(7.67)**	(13.72)**
FC*Iadj	**0.152**	**-0.107**	**0.088**	**0.592**	**0.605**	**0.158**	**0.716**	**1.95**
	(4.58)**	**(-3.28)****	**(3.64)****	**(5.76)****	**(2.90)****	**(0.86)**	**(5.04)****	**(3.24)****
FC	-0.019	0.042	0.017	0.095	0.036	0.024	0.068	0.142
	(-5.29)**	(20.02)**	(5.90)**	(11.01)**	(3.62)**	(4.03)**	(7.94)**	(4.28)**
Size	0.001	0.007	0.01	0.009	-0.023	-0.029	-0.004	-0.022
	(1.08)	(9.78)**	(8.27)**	(11.27)**	(-10.10)**	(-15.98)**	(-1.26)	(-9.47)**
M2B	-0.022	-0.025	-0.028	-0.027	-0.01	-0.008	-0.022	-0.012
	(-28.33)**	(-32.51)**	(-29.80)**	(-33.83)**	(-2.56)*	(-2.10)*	(-5.03)**	(-2.91)**
LR					-0.202	-0.215	-0.214	-0.205
					(-11.16)**	(-11.74)**	(-11.60)**	(-11.22)**
CF	-0.134	-0.123	-0.126	-0.121	-0.125	-0.131	-0.109	-0.113
	(-22.38)**	(-20.93)**	(-20.96)**	(-20.06)**	(-3.89)**	(-4.05)**	(-3.43)**	(-3.49)**
NBER	0.000025	0.00039	0.00031	0.0012	-0.045	-0.044	-0.042	-0.042
	(0.02)	(0.24)	(0.19)	(0.74)	(-6.91)**	(-6.75)**	(-6.38)**	(-6.46)**
DefSpr	-1.639	-1.501	-1.703	-1.567	4.134	4.58	4.307	4.536
	(-4.09)**	(-3.79)**	(-4.26)**	(-3.92)**	(5.78)**	(6.35)**	(6.02)**	(6.28)**
r_f	-0.118	-0.034	-0.102	-0.069	0.324	0.297	0.211	0.298
	(-2.45)*	(-0.71)	(-2.11)*	(-1.43)	(2.43)*	(2.19)*	(1.58)	(2.23)*
Intercept	0.1	0.034	0.104	0.066	0.225	0.252	0.372	0.255
	(13.22)**	(5.34)**	(13.74)**	(10.51)**	(11.84)**	(14.07)**	(17.14)**	(14.89)**
N	44,858	44,858	44,858	44,858	51,875	51,875	51,875	51,875

* $p < 0.05$; ** $p < 0.01$

Table 1.8: Regressions of Investment on Expected Volatility and its Interaction with Financial Constraints Index with Control of Cash Holdings

This table presents results from the regressions of investment on expected volatility and its interaction with financial constraints index with control of cash holdings. The four columns *Size*, *Payout*, *SA* and *WW* shows the four financial constraints classification schemes which are described in detail in *Table 1.5*.The dependent variable is the industry-adjusted investment intensity I_{adj} at the beginning of the year. The independent variables are the expected industry volatility σ_{Ind} at the beginning of the year, the financial constraints index *FC*, the Altman's *Z-Score* and the interaction of σ_{Ind} with *FC*. *FC* is a dummy variable which equals to 1 if the firm is financially constrained and 0 otherwise for the *Size* and *Payout* scheme whereas it equals *SA* index or *WW* index for the other two schemes. *Cashholdings* is total cash over beginning-of-period total asset. The other control variables are defined in *Table 1.3*. The sample data is taken from CRSP and COMPUSTAT for the period of 1965 to 2009. Only non-financial and non-utility firms are included. The full sample contains 7,680 firms. The regression coefficients and t-statistics are calculated based on Newey-west standard errors with 6 lags.

	Size	**Payout**	**SA**	**WW**
σ_{Ind}	0.013	0.017	-0.065	-0.018
	(1.51)	(1.76)	(-2.83)**	(-2.72)**
FC	0.01	0.019	0.014	0.011
	(2.63)**	(7.53)**	(8.31)**	(1.61)
σ_{Ind}***FC**	**-0.036**	**-0.031**	**-0.017**	**-0.112**
	(-2.82)**	**(-2.76)****	**(-2.56)***	**(-3.87)****
Cashholdings	0.00055	-0.00044	-0.002	-0.00027
	(0.58)	(-0.49)	(-1.85)	(-0.32)
Size	0.00009	0.001	0.003	-0.001
	(0.18)	(2.37)*	(8.49)**	(-3.75)**
M2B	0.008	0.009	0.007	0.009
	(15.58)**	(21.11)**	(17.31)**	(23.25)**
LR	-0.021	-0.025	-0.019	-0.018
	(-6.76)**	(-9.15)**	(-7.61)**	(-7.40)**
CF	0.053	0.057	0.061	0.058
	(16.47)**	(20.21)**	(22.52)**	(20.85)**
NBER	-0.002	-0.003	-0.002	-0.003
	(-2.26)*	(-3.30)**	(-3.15)**	(-4.04)**
DefSpr	0.35	0.319	0.225	0.193
	(3.46)**	(3.39)**	(2.93)**	(2.49)*
r_f	0.134	0.14	0.104	0.114
	(6.29)**	(7.26)**	(6.36)**	(6.91)**
Intercept	-0.016	-0.027	0.023	-0.003
	(-3.30)**	(-8.58)**	(4.13)**	(-1.23)
N	32,514	42,038	54,266	54,266

* $p < 0.05$; ** $p < 0.01$

Figure 1.1: The Optimal Firm and Asset Value over a Set of Financial Constraints Costs for Different Debt Maturities

This figure shows the optimal firm values and the decision to give up risk-shifting graphically over a set of financial constraints costs. The total financial constraints cost M is defined as $M = M_0(1+i)$ where the fixed cost $M_0 = 30$ and the percentage financial constraints cost $i \in [0, 0.55]$. The upper panel shows the pattern of the initial firm value with respect to M for debt rollover rates $m = 0, 0.1 and 0.2$ which corresponds to debt maturities of ∞, 10 years and 5 years. The lower panel shows the equity values evaluated with the risk-shifting model ($SL(C)$) and the benchmark model ($SB(C)$) for the 10-year debt scenario. The firm has an initial cash flow $X_0 = 100$. The marginal tax rate $\tau = 20\%$ and the expected rate of return $\mu = 1\%$. The pre- and post-shifting risk levels are $\sigma_L = 25\%$ and $\sigma_H = 40\%$. The percentage cost of bankruptcy is $\alpha = 0.25$ and the risk-free rate is $r_f = 6\%$.

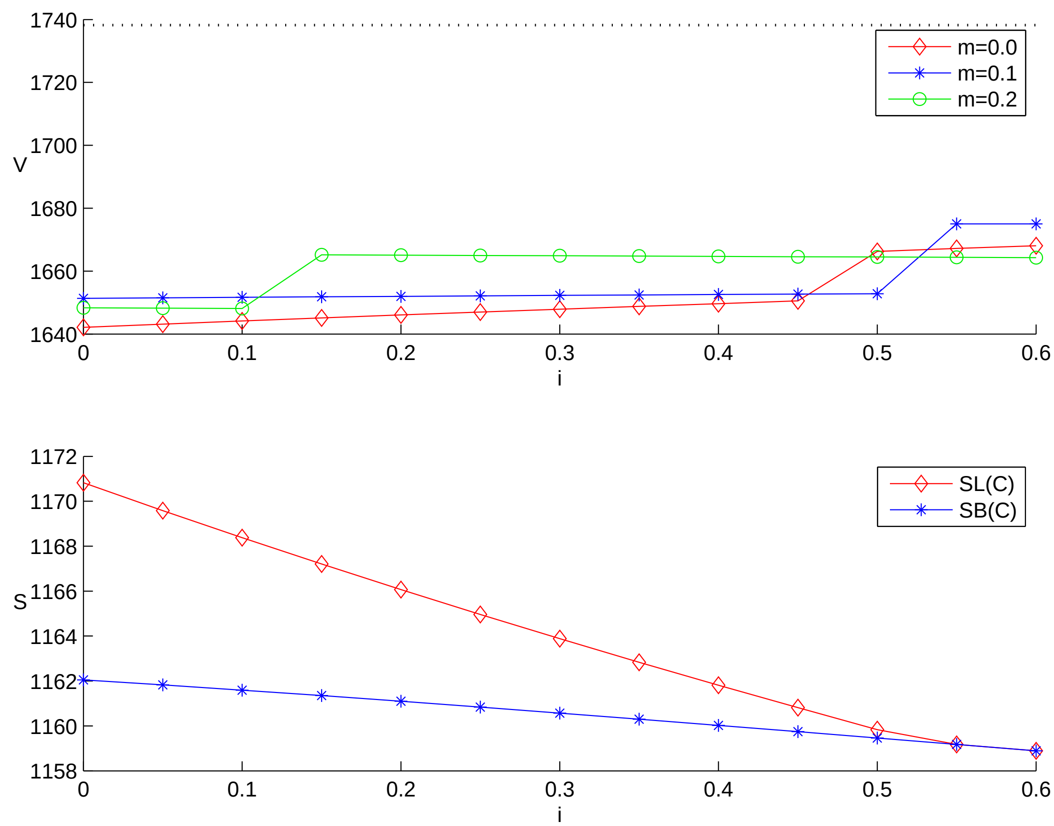


Figure 1.2: The Optimal Risk-shifting Incentives v.s. Financial Constraints Costs for Different Debt Maturities

This figure shows the optimal absolute and relative risk-shifting incentives graphically over a set of financial constraints costs and debt maturities. The total financial constraints cost M is defined as $M = M_0(1+i)$ where the fixed cost $M_0 = 30$ and the percentage financial constraints cost $i \in [0, 0.55]$. The debt rollover rates $m \in [0, 0.2]$ corresponds to debt maturities of ∞ down to 5 years ($T = \frac{1}{m}$). The left panel shows the pattern of the optimal risk-shifting threshold XR, and the right panel shows the relative measure of risk-shifting incentive: $K = \frac{XR}{XDL}$ where XDL is the default threshold should the firm gives up risk-shifting. The firm has an initial cash flow $X_0 = 100$. The marginal tax rate $\tau = 20\%$ and the expected rate of return $\mu = 1\%$. The pre- and post-shifting risk levels are $\sigma_L = 25\%$ and $\sigma_H = 40\%$. The percentage cost of bankruptcy is $\alpha = 0.25$ and the risk-free rate is $r_f = 6\%$.

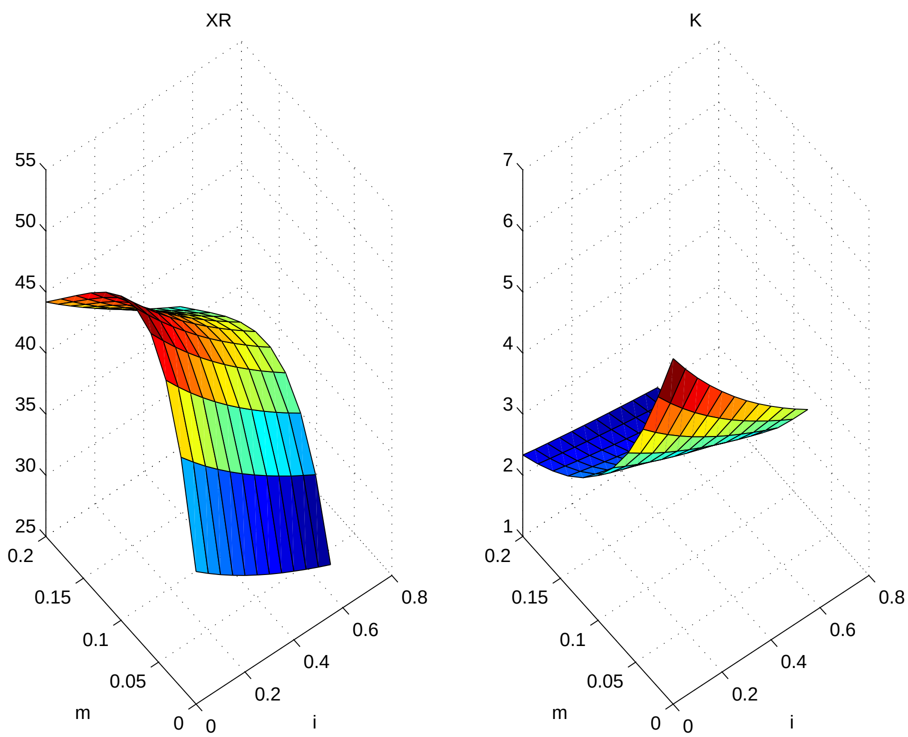
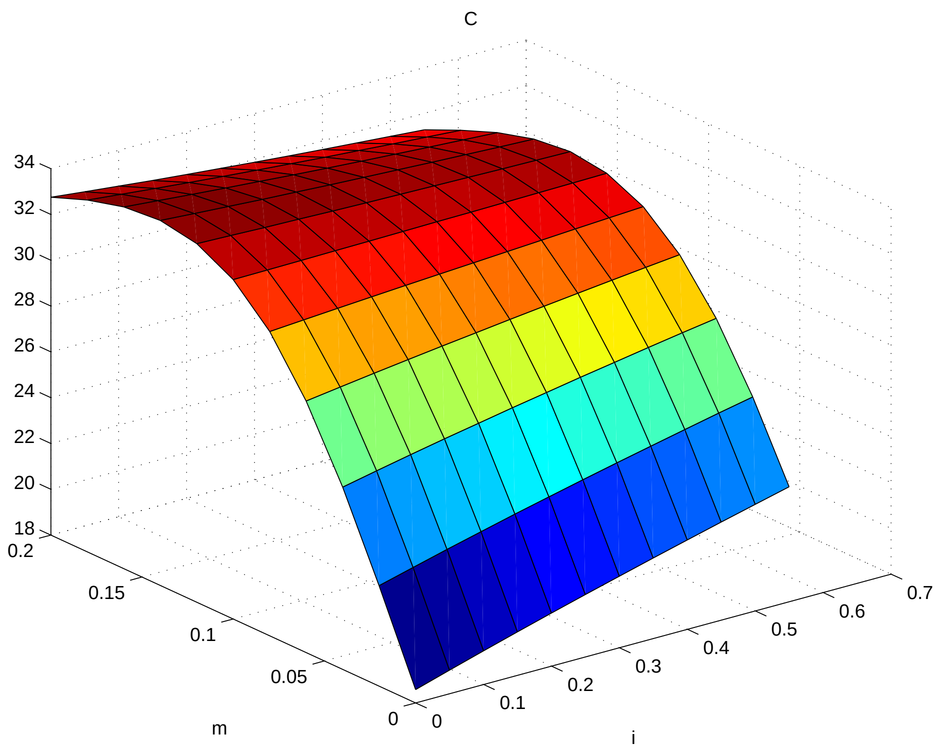


Figure 1.3: The Optimal Coupon v.s. Financial Constraints Costs for Different Debt Maturities

This figure shows the optimal coupon C graphically over a set of financial constraints costs and debt maturities. The total financial constraints cost M is defined as $M = M_0(1 + i)$ where the fixed cost $M_0 = 30$ and the percentage financial constraints cost $i \in [0, 0.55]$. The debt rollover rates $m \in [0, 0.2]$ corresponds to debt maturities of ∞ down to 5 years ($T = \frac{1}{m}$). The firm has an initial cash flow $X_0 = 100$. The marginal tax rate $\tau = 20\%$ and the expected rate of return $\mu = 1\%$. The pre- and post-shifting risk levels are $\sigma_L = 25\%$ and $\sigma_H = 40\%$. The percentage cost of bankruptcy is $\alpha = 0.25$ and the risk-free rate is $r_f = 6\%$.



Chapter 2

Dynamic Investment and Financing

2.1 Introduction

Since Modigliani and Miller (1958), the determinants of capital structure have been a central topic of corporate finance research. Researchers have considered trading off tax benefit of debt against bankruptcy costs, information asymmetries leading to a pecking order of internal and external financing, market timing to benefit from market inefficiencies, several forms of agency costs, managerial inertia, and other imperfections. Empirical tests have identified determinants of leverage ratios to distinguish empirical predictions of different theories. Arguably, the trade-off theory receives most empirical support, yet it is still challenged by some empirical findings which favor other theories. Therefore, there is no clear consensus in the literature. Furthermore, none of the extant theories answers the following questions in a unified framework: (1) why firms tend to use debt financing so conservatively, (2) whether there is a target leverage ratio and partial adjustment towards it, and (3) why average leverage paths persist for over two decades. This paper attempts to fill this void by exploring a dynamic framework of corporate investment and financing decisions.

Regarding the first question, Graham (2000) finds that firms, even stable and profitable, use less debt than predicted by the static view of tax benefit of debt. Two out of five firms have an average leverage ratio of less than 20%, and the median firm uses only 31.4% leverage over the 1965 to 2000 period, which creates a "low leverage puzzle." In addition, Strebulaev and Yang (2012) find that on average 10% of firms have zero leverage and almost 22% firms have less than 5% quasi-market leverage, which represents a "zero leverage puzzle." Extending the static trade-off model into a dynamic one with financing costs helps to significantly reduce the model predicted market leverage to 36% on average in the cross section and 26% at refinancing points in Strebulaev (2007). However, these model simulated leverage ratios are still higher than those observed empirically. We emphasize the importance of endogenous investment in a dynamic trade-off model and also provide an economic mechanism that explains why firms tend to use debt financing so conservatively. Our model simulated data sets feature, on average, 20% leverage in dynamics and 19% at investment points.

The second issue is related to the trade-off and the partial adjustment theories. An important implication of trade-off theory is that firms have fixed target (i.e. optimal) leverage ratios. Whenever actual leverage deviates from target leverage, the firm makes adjustments. This hypothesis is supported by Fama and French (2002), among others, who find that firms' leverage ratios adjust, albeit slowly, to their targets. Flannery and Rangan (2006) find that, on average, the speed of adjustment is 30% of the gap per year. In contrast, Baker and Wurgler (2002) argue that managers actively exploit stock market mispricing and issue stocks when they are over priced, so leverage ratios reflect such attempts in the past. Welch (2004), on the other hand, claims that managerial inertia leaves

[1]This chapter is related to the working paper "Dynamic Investment and Financing," with Dirk Hackbarth.

stock price changes to dominantly affect leverage ratios. In these two articles, the effects on leverage are persistent so that there is no target leverage and no partial adjustment. Notably, Leary and Roberts (2005) defend the partial adjustment findings by arguing that the persistence identified by the previous two studies is largely due to the inaction induced by the presence of adjustment cost. Furthermore, Tserlukevich (2008) provides theoretical evidence on gradual and lumpy leverage adjustments by considering the effect of real frictions instead of financing frictions. Similarly, we establish that dynamic trade-off models without endogenous investment (1) overestimate target leverage ratios, and (2) can be misleading in that their target leverage ratios ignore the investment process. The latter point means that there is no meaningful measurement of partial adjustment to target leverage without recognition of the structure of the investment process.

The third question relates to Lemmon, Roberts and Zender (2008), who document persistence of average leverage of four portfolios, which are based on the level the leverage, for very long periods of time. That is, they find that leverage ratios tend to converge to medium levels over time, but high (low) leverage firms preserve their high (low) leverage ratios for over twenty years. They conclude that their findings are puzzling, because they cannot be explained by previously identified capital structure determinants or models. It is therefore remarkable that our dynamic model can generate average leverage ratios that are path dependent and persistent for very long periods of time.

In essence, real frictions in a dynamic trade-off model can generate these three regularities and hence can go a long way in explaining capital structure. More specifically, the insights of this paper derive from the simple observation that companies make sequential investments. These investments are a result of breaking down projects into multiple stages, and the implementation of one stage is contingent on the completion of the previous stages. Hence financing decisions of investments are sequential too, so the structure of the investment process strongly affects financing decisions.

More specifically, we consider a two-stage investment option, which is financed by an optimal mix of equity and debt in each stage.[2] The two-stage investment option has two scales, Π_1 and Π_2, which determine the *structure of the investment process*. If the second stage is more profitable than the first stage, then we refer to it as a "back-loaded" investment process. The optimal leverage ratios for both stages along with the endogenous investment and default thresholds maximize equity value. The model's solution features an *intertemporal effect*: firms use their debt capacity to balance the benefits from accelerating the first option with increased debt financing against the costs of delaying the second option due to decreased debt financing. Conditional on their investment and financing opportunities, juvenile firms underutilize debt when financing investment the first time to retain financial flexibility. Underutilization of debt persists when adolescent firms mature (i.e. exercise their last investment options), and it is more (less) severe for more back-loaded (front-loaded) investment opportunities. Thus, leverage dynamics crucially hinge upon the structure of the investment process and otherwise identical firms appear to have significantly different target leverage ratios.

We estimate the key model parameters via Simulated Method of Moments (SMM). In particular, structural estimation determines the two investment scales, bankruptcy cost, and effective tax rate. Intuitively, SMM finds the set of parameters, which minimizes the difference of the simulated model moments and the data moments from COMPUSTAT's annual tapes for the period of 1965 to 2009. We then split the full sample into low, medium, and high market-to-book (or Q) subsamples, and employ SMM also to fit the four parameters for each subsample. We split the sample based on Q, because it is the most reliable proxy for growth opportunities (see, e.g., Adam and Goyal (2007)).

[2] Although we study two stages, our model can be extended to multiple stages. Because its implications will be qualitatively identical, we refer to it as multi-stage model and compare it to an otherwise identical single-stage model.

Low Q firms tend to have fewer investment opportunities, whereas high Q firms tend to have more investment opportunities. Therefore, the relative value of Q is informative about the structure of the investment process in the real data. Indeed, our estimation results reveal that high Q firms have the most back-loaded investment processes, and low Q firms have the most front-loaded ones.

Using the structural estimation results for the full sample, we perform capital structure regressions on model simulated data and show that our model can replicate stylized facts established by empirical research. In the spirit of Strebulaev (2007), simulation of the multi-stage model of corporate investment and financing dynamics reinforces the need to differentiate investment points from other data points when interpreting coefficient estimates for market-to-book or profitability in a dynamic world. Moreover, we find that leverage is negatively related to the risk of cash flows, the cost of bankruptcy, and market-to-book, but positively related the size of the firm and the tax rate, which are all consistent with the literature (see, e.g., Frank and Goyal (2009)).

Finally, we document that real frictions in a dynamic model can produce average leverage paths that closely resemble the corresponding ones documented by Lemmon, Roberts and Zender (2008). Put differently, optimal investment and financing decisions in a dynamic model can largely explain the otherwise puzzling patterns that, despite of some convergence, average leverage ratios across the four portfolios are fairly stable over time for both types of sorts (i.e. actual and unexpected leverage) performed by these authors.[3] In particular, an extension of our dynamic framework to randomly imposed initial variation in leverage reveals that, if model firms are "born" with high (low) leverage ratios, then they maintain their relatively high (low) levels for over 20 years in spite of the fact that leverage ratios converge somewhat to more moderate levels over time. This result illustrates that firms, which know the structure of their investment processes, look very far into the future and make decisions on debt usage accordingly. This leads to fairly stable leverage ratios, and serves in the simulations as an important, unobserved determinant of the permanent component of leverage.

This paper builds on dynamic capital structure models by Fischer, Heinkel, and Zechner (1989) and Goldstein, Ju and Leland (2001), with the latter providing the contingent claims valuation framework adopted in this paper. Similarly, Strebulaev (2007) studies a dynamic trade-off model with optimal refinancing in the presence of adjustment costs, which are constant over time. We examine whether his conclusions carry over to our simulated economies with real frictions that vary over time. Like us, Hennessy and Whited (2005) analyze a dynamic trade-off model with investment and perform structural estimation. They largely focus on role of tax regimes, while we focus on the intertemporal effect. More recently, Tserlukevich (2008) also invokes real fictions to produce gradual and lumpy leverage adjustments in the absence of financial frictions. Two key differences are that there is no intertemporal effect in his model and that leverage ratios produced by his model are much higher than ours. Finally, DeAngelo, DeAngelo, and Whited (2011) study transitory debt that arises due to unexpected (positive) shocks to investment opportunities whose properties are uncertain as of time zero. Our model firm optimizes at time zero knowing the structure of the investment process, so there is no role for transitory debt and yet we obtain average leverage persistence results similar to DeAngelo, DeAngelo, and Whited (2009).

The rest of the paper is organized as follows. Section 2.2 presents and solves the model. Section 2.3 studies the intertemporal effect. Section 2.4 estimates structural parameters using SMM. Section 2.5 presents the results of cross-sectional capital structure regressions with model simulated data. Section 2.6 examines capital structure persistence with model simulated data. Section 2.7 concludes.

[3] In this part of the paper, we employ the structural estimation results for the three subsamples to introduce industry variation, so sorting on "unexpected leverage" defined as the residuals from a cross-sectional regression of leverage on firm characteristics and industry indicator variables is different from sorting on actual leverage.

2.2 Model

This section provides a simple framework to analyze the dynamics of corporate investment and financing. In particular, we consider two versions of a model with endogenous financing and investment decisions. While the multi-stage model features two sequentially exercisable investment options, where the capital expenditure in each stage is financed by a mixture of equity and debt, the single-stage model has only one investment option. The single-stage model serves as a benchmark to gauge interactions between investment and financing in the otherwise identical multi-stage model.

2.2.1 Setup

We consider a partial equilibrium model of corporate investment and financing dynamics. Time t is continuous and uncertainty is modeled by a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Corporate assets generate a continuous stream of cash flows, X_t, which evolve for $t > 0$ according to a geometric Brownian motion with drift μ, volatility σ, and initial value initial cash flow $X_0 > 0$ at time $t = 0$. Corporate taxes are paid on cash flows at a constant rate τ based on full loss offset provisions. Agents are risk-neutral and discount cash flows at a constant interest rate $r > \mu$.

At time $t = 0$, the firm has no assets-in-place and a two-stage project, i.e. a compound option in that the implementation of the second stage investment is contingent upon the completion of the first stage. The two constants Π_1 and Π_2 represent the scales of the two investment options. Suppressing time dependence of cash flows, $\Pi_i X$ is the cash flow from investing in stage $i = 1, 2$, which requires a capital expenditure, F_i.

The investment cost, F_i, can be financed with a mix of debt and equity. We assume that debt has an infinite maturity and denote the coupon rate on debt issued in stage i by C_i.[4] The optimal time to invest is the one that maximizes the market value of equity. The optimal time to default on debt coupon payments is also endogenously determined (i.e. maximizes equity value). In the event of default, equityholders receive nothing and debtholders assume ownership of the firm's assets net of bankruptcy costs. Bankruptcy costs include the loss of interest tax shields, the loss of the second-stage option (if it has not been exercised), and the fraction α of the value of assets-in-place. The endogenous investment thresholds and default thresholds in stage i are X_{Si} and X_{Di} for $i = 1, 2$.[5]

At time $t = 0$, equityholders wait to exercise the first investment option, which is triggered when cash flows rise to the investment threshold $X_{S1} \in (X_0, \infty)$ for the first time from below. We denote this waiting period "stage 0" and call the firm in this stage a "juvenile" firm. When the first option is exercised, the firm issues debt D_1 and equity E_1 to finance the fixed investment cost F_1. Then the firm enters "stage 1" and becomes an "adolescent" firm.

In stage 1, the firm has assets-in-place, another investment option, and a default option because of D_1. If cash flows decline to the default threshold $X_{D1} \in (0, X_{S1})$ before reaching the second investment threshold $X_{S2} \in (X_{S1}, \infty)$, equityholders default. On the other hand, if cash flows reach the investment threshold X_{S2} before decreasing to X_{D1}, equityholders exercise the second option and finance the investment cost F_2 with a mix of debt, D_{22}, and equity, E_2. We assume that D_{22} has the same seniority as D_1 whose value is denoted as D_{21} in the second stage. The firm then enters "stage 2" and becomes a "mature" firm.

In stage 2, the firm has assets-in-place, no further investment options, and a default option because of D_{21} and D_{22}. Equityholders default when X touches the default threshold $X_{D2} \in (0, X_{S2})$

[4] The reliance on consol bonds simplifies the analysis substantially but does not alter the economic insights.

[5] See Appendix A and Section 2.6 for a multi-stage model extended to have also an initial debt coupon C_0.

for the first time from above. Finally, we assume that there are no conflicts of interest and the debt coupons maximize equity value at time 0. Table 2.1 provides a notational key.

To quantify the dynamic interactions of endogenous investment and financing decisions, we use a single-stage model as a benchmark. This single-stage model is a truncated version of the multi-stage model in that there is no intermediate stage 1. In stage 0, the juvenile firm has no asset-in-place and no debt. It makes one investment of scale Π when X touches the investment threshold X_S from below for the first time, and then becomes a mature firm (i.e. enters stage 2). The capital expenditure, F, is funded by a mixture of debt and equity, where C denotes the coupon of the firm's debt issue.

2.2.2 Solution of the Multi-Stage Model

Using backward induction, we obtain values of debt and equity in each stage and then pin down the optimal investment and default thresholds via smooth-pasting conditions.

Mature Firm (Stage 2)

In the second stage, both investment options have been converted into assets-in-place. The new debt, D_{22}, is issued to partially finance the investment cost F_2 and the debt issued in the first stage, D_{21}, offer tax savings but give rise to a default decision (or threshold X_{D2}). Following standard arguments, the debt value $D_{2i}(X, C_1, C_2)$, $i = 1, 2$, has for $X \geq X_{D2}$ a solution of the form:[6]

$$D_{2i}(X, C_1, C_2) \;=\; A_{1i}X^a + A_{2i}X^z + \frac{C_i}{r}, \tag{2.1}$$

where the two exponents $a < 0$ and $z > 1$ are the negative and positive roots of the quadratic equation $y(y-1)\sigma^2/2 + y\mu - r = 0$. The constants A_{1i} and A_{2i} solve the following boundary conditions. When $X \uparrow \infty$, debt becomes risk-free and its value equals the present value of a perpetuity: $D_{2i}(\infty, C_1, C_2) = \frac{C_i}{r}$. On the other hand, when X declines to X_{D2}, equityholders default and the owners of D_{2i} get a proportion of the liquidation value based on the coupon C_i for $i = 1, 2$:[7]

$$D_{2i}(X_{D2}, C_1, C_2) \;=\; \frac{C_i}{C_1 + C_2}\frac{(1-\alpha)(1-\tau)(\Pi_1 + \Pi_2)X_{D2}}{r-\mu}, \; i = 1, 2. \tag{2.2}$$

Equity value $E_2(X, C_1, C_2)$, on the other hand, has for $X \geq X_{D2}$ a solution of the form:

$$E_2(X, C_1, C_2) \;=\; B_1X^a + B_2X^z + (1-\tau)\Big(\frac{(\Pi_1 + \Pi_2)X}{r-\mu} - \frac{(C_1 + C_2)}{r}\Big), \tag{2.3}$$

where the constants B_1 and B_2 satisfy the following boundary conditions:

$$E_2(X_{D2}, C_1, C_2) \;=\; 0, \tag{2.4}$$

$$E_2(\infty, C_1, C_2) \;=\; (1-\tau)\Big(\frac{(\Pi_1 + \Pi_2)X}{r-\mu} - \frac{(C_1 + C_2)}{r}\Big). \tag{2.5}$$

Simple algebra yields the value of the two debt issues for $X \geq X_{D2}$:

$$D_{2i}(X, C_1, C_2) \;=\; \frac{C_i}{r}\Big(1 - (\frac{X}{X_{D2}})^a\Big) + \frac{C_i}{C_1 + C_2}\frac{(1-\alpha)(1-\tau)(\Pi_1 + \Pi_2)X_{D2}}{r-\mu}(\frac{X}{X_{D2}})^a, \tag{2.6}$$

[6] See, e.g., Goldstein, Ju, and Leland (2001) and Hackbarth, Hennessy, and Leland (2007) for details.
[7] We use equal priority for D_{2i}. See, e.g., Hackbarth and Mauer (2012) for other priority structures.

with $i = 1, 2$, and also the value of equity for $X \geq X_{D2}$:

$$E_2(X, C_1, C_2) = (1-\tau)\Big(\frac{(\Pi_1+\Pi_2)X}{r-\mu} - \frac{C_1+C_2}{r} - (\frac{(\Pi_1+\Pi_2)X_{D2}}{r-\mu} - \frac{C_1+C_2}{r})(\frac{X}{X_{D2}})^a\Big), \quad (2.7)$$

where $(\frac{X}{X_{D2}})^a$ denotes the state price for default. Finally, the total firm value in this stage is the sum of D_{21}, D_{22} and E_2.

As mentioned earlier, the only decision that equityholders makes in this stage is when to default. The optimal default threshold, X_{D2}, is the one that maximizes the value of equity, E_2:

$$\frac{\partial E_2(X, C_1, C_2)}{\partial X}\Big|_{X=X_{D2}} = 0, \quad (2.8)$$

which yields a closed-form solution for the endogenous default threshold in the second stage:

$$X_{D2} = \frac{a\,(C_1+C_2)\,(r-\mu)}{r\,(a-1)\,(\Pi_1+\Pi_2)}. \quad (2.9)$$

Adolescent Firm (Stage 1)

In the first stage, only the second investment option has not yet been exercised. The adolescent firm has assets-in-place and its capital structure consists of a mixture of debt, D_1, and equity, E_1. It has both an option to default and an option to invest, so it solves a joint financing and investment problem. When cash flows rise to the investment threshold, X_{S2}, equityholders exercise the second option and issues D_{22}. Equityholders default if cash flows decline to the default threshold, X_{D1}.

The values of debt, D_1, and equity, E_1, have solutions similar to the ones in equation (2.1) and (2.3), but obey different boundary conditions. When $X \downarrow X_{D1}$, debtholders receive the liquidation value: $D_1(X_{D1}, C_1, C_2) = \frac{(1-\alpha)(1-\tau)\Pi_1 X_{D1}}{r-\mu}$. If the firm keeps growing and X rises to the investment threshold, X_{S2}, equityholders exercise the second-stage investment option, and debt value from stage 1 equals debt value in stage 2: $D_1(X_{S2}, C_1, C_2) = D_{21}(X_{S2}, C_1, C_2)$. For $X_{D1} \leq X \leq X_{S2}$, these value-matching conditions imply the following solution for value of debt in stage 1:

$$\begin{aligned} D_1(X, C_1, C_2) &= \frac{C_1}{r}\Big(1 - L(X) - (\frac{X_{S2}}{X_{D2}})^a H(X)\Big) + (1-\alpha)(1-\tau)\Big(\frac{\Pi_1 X_{D1}}{r-\mu}L(X) + \\ &\quad \frac{C_1}{C_1+C_2}\frac{(\Pi_1+\Pi_2)X_{D2}}{r-\mu}(\frac{X_{S2}}{X_{D2}})^a H(X)\Big), \end{aligned} \quad (2.10)$$

where

$$L(X) = \frac{X^z X_{S2}^a - X^a X_{S2}^z}{X_{D1}^z X_{S2}^a - X_{D1}^a X_{S2}^z} \quad (2.11)$$

$$H(X) = \frac{X_{D1}^z X^a - X_{D1}^a X^z}{X_{D1}^z X_{S2}^a - X_{D1}^a X_{S2}^z} \quad (2.12)$$

denote state prices for default and investment. Intuitively, debt value, D_1, is the weighted average of the present value of the coupon payments C_1 up until default in either the first or the second stage, and the liquidation value that debtholders get when equityholders default in either stage.

Equity value E_1, on the other hand, approaches zero when $X \downarrow X_{D1}$: $E_1(X_{D1}, C_1, C_2) = 0$. As $X \uparrow X_{S2}$, it satisfies $E_1(X_{S2}, C_1, C_2) = E_2(X_{S2}, C_1, C_2) - [F_2 - D_{22}(X_{S2}, C_1, C_2)]$, because the fixed investment cost, F_2, is funded by a mixture of debt and equity. For $X_{D1} \leq X \leq X_{S2}$, these value-matching conditions yield the following solution for the value of equity in stage 1:

$$
\begin{aligned}
E_1(X, C_1, C_2) \;=\; & (1-\tau)\Big[\big(\frac{\Pi_1 X}{r-\mu} - \frac{C_1}{r}\big) - \big(\frac{\Pi_1 X_{D1}}{r-\mu} - \frac{C_1}{r}\big)L(X) + \Big(\frac{\Pi_2 X_{S2}}{r-\mu} - \frac{C_2}{r} - \\
& \frac{F_2 - D_{22}(X_{S2}, C_1, C_2)}{1-\tau} - \big(\frac{(\Pi_1+\Pi_2)X_{D2}}{r-\mu} - \frac{C_1+C_2}{r}\big)\big(\frac{X_{S2}}{X_{D2}}\big)^a\Big)H(X)\Big].
\end{aligned} \tag{2.13}
$$

The first two terms in square brackets of equation (2.13) denote the present value of after-tax cash flows to equityholders until equityholders default in the current stage. The next few terms in this equation show the value from entering the next stage. Given the value of E_1, equityholders determine the optimal default threshold, X_{D1}, by maximizing equity value:

$$
\frac{\partial E_1(X, C_1, C_2)}{\partial X}\Big|_{X=X_{D1}} = 0. \tag{2.14}
$$

Similarly, the optimal investment threshold, X_{S2}, solves the smooth-pasting condition:

$$
\frac{\partial E_1(X, C_1, C_2)}{\partial X}\Big|_{X=X_{S2}} = \frac{\partial E_2(X, C_1, C_2)}{\partial X}\Big|_{X=X_{S2}} + \frac{\partial D_{22}(X, C_1, C_2)}{\partial X}\Big|_{X=X_{S2}}. \tag{2.15}
$$

Juvenile Firm (Stage 0)

In the initial stage, the juvenile firm has no assets-in-place, an option on a two-stage investment project, and no pre-existing debt.[8] The value of equity in this stage, E_0, has a solution similar to the one in equation (2.3) but without the last term on the right-hand side. As $X \downarrow 0$, equity value goes to zero: $E_0(0, C_1, C_2) = 0$. When X touches the investment threshold X_{S1} for the first time from below, the option is exercised, and debt and equity finance the exercise cost, F_1: $E_0(X_{S1}, C_1, C_2) = E_1(X_{S1}, C_1, C_2) - [F_1 - D_1(X_{S1}, C_1, C_2)]$. For $X \leq X_{S1}$, this yields the following solution:

$$
\begin{aligned}
E_0(X, C_1, C_2) \;=\; & (1-\tau)\big(\frac{X}{X_{S1}}\big)^z\Big[\big(\frac{\Pi_1 X_{S1}}{r-\mu} - \frac{C_1}{r}\big) - \frac{F_1 - D_1(X_{S1}, C_1, C_2)}{1-\tau} - \\
& \big(\frac{\Pi_1 X_{D1}}{r-\mu} - \frac{C_1}{r}\big)L(X) + \Big(\frac{\Pi_2 X_{S2}}{r-\mu} - \frac{C_2}{r} - \frac{F_2 - D_{22}(X_{S2}, C_1, C_2)}{1-\tau} - \\
& \big(\frac{(\Pi_1+\Pi_2)X_{D2}}{r-\mu} - \frac{C_1+C_2}{r}\big)\big(\frac{X_{S2}}{X_{D2}}\big)^a\Big)H(X)\Big].
\end{aligned} \tag{2.16}
$$

Equity value in this stage equals the present value of after-tax cash flows conditional on exercise of the first-stage option until equityholders default in stage 1 (the first three terms in square brackets of equation (2.16)). If cash flows grow further and the firm expands a second time, then it enters into stage 2 with the added value given by the next few terms in square brackets of equation (2.16).

In this stage, equityholders' choose when to invest and how much debt and equity to issue to finance the investment cost, F_1. When X is low, the benefit from exercising the option is outweighed by the value of waiting-to-invest, hence the equityholders keep the option alive. When X rises sufficiently, equityholders exercise the first option at X_{S1}, which solves the smooth-pasting condition:

$$
\frac{\partial E_0(X, C_1, C_2)}{\partial X}\Big|_{X=X_{S1}} = \frac{\partial E_1(X, C_1, C_2)}{\partial X}\Big|_{X=X_{S1}} + \frac{\partial D_1(X, C_1, C_2)}{\partial X}\Big|_{X=X_{S1}}. \tag{2.17}
$$

Finally, the debt coupons C_1 and C_2 maximize initial equity value, $E_0(X_0, C_1, C_2)$, subject to the above-mentioned smooth-pasting conditions for the thresholds X_{S1}, X_{S2}, X_{D1} and X_{D2}.

[8] We relax this assumption in Appendix A and Section 2.6 to study the effect an initial debt coupon C_0.

2.2.3 Solution of the Single-Stage Model

Claim values and optimal exercise thresholds for this truncated version of the model are also solved recursively. When $X \uparrow \infty$ in the last stage, the firm becomes risk-less, so its equity value $E_{B1}(X, C)$ and its debt value $D_{B1}(X, C)$ converge to the present value of perpetual dividend and interest payments. On the other hand, when $X \downarrow X_D$, equityholders default, and debtholders obtain the entire liquidation value (net of bankruptcy costs).

In the inital stage, the firm has a single investment option, so firm value equals the value of this option. If X decreases to zero, the option becomes worthless: $E_{B0}(0, C) = 0$. But if X rises to the investment threshold, X_S, the option is exercised: $E_{B0}(X_S, C) = E_{B1}(X_S, C) - [F - D_{B1}(X_S, C)]$.

The above-mentioned boundary conditions yield the following solutions for $X \geq X_D$:

$$D_{B1}(X,C) = \frac{C}{r}\Big[1-(\frac{X}{X_D})^a\Big] + (1-\alpha)(1-\tau)\frac{\Pi X_D}{r-\mu}(\frac{X}{X_D})^a, \tag{2.18}$$

$$E_{B1}(X,C) = (1-\tau)\Big[\frac{\Pi X}{r-\mu} - \frac{C}{r} - (\frac{\Pi X_D}{r-\mu} - \frac{C}{r})(\frac{X}{X_D})^a\Big], \tag{2.19}$$

and for $X \leq X_S$:

$$E_{B0}(X,C) = \Big[(1-\tau)\frac{\Pi X_S}{r-\mu} + \frac{\tau C}{r} - F - [(1-\tau)\frac{\Pi \alpha X_D}{r-\mu} + \frac{\tau C}{r}](\frac{X_S}{X_D})^a\Big](\frac{X}{X_S})^z. \tag{2.20}$$

The optimal default threshold, X_D, maximizes the value of equity, E_{B1}, that is:

$$\frac{\partial E_{B1}(X,C)}{\partial X}\Big|_{X=X_D} = 0, \tag{2.21}$$

which implies the following closed-form solution:

$$X_D = \frac{a\,C\,(r-\mu)}{(a-1)\,\Pi\, r}. \tag{2.22}$$

The optimal investment threshold, X_S, also maximizes the value of equity, E_{B1}, that is:

$$\frac{\partial E_{B0}(X,C)}{\partial X}\Big|_{X=X_S} = \frac{\partial E_{B1}(X,C)}{\partial X}\Big|_{X=X_S} + \frac{\partial D_{B1}(X,C)}{\partial X}\Big|_{X=X_S}. \tag{2.23}$$

Finally, debt coupon C maximizes in initial equity value, E_{B0}, subject to the above-mentioned smooth-pasting conditions for the thresholds X_D and X_S.

2.3 Financial Flexibility and the Investment Process

In this section, we study the key difference between the multi-stage model and the single-stage model, which is the intermediate stage of an adolescent firm. This stage creates an intertemporal effect, which links financial flexibility to the investment process. We illustrate how financing and investment decisions of the adolescent firm influence those of the mature firm and vice versa.

2.3.1 Intertemporal Effect

Consider the firm in the multi-stage model at time zero. On the one hand, suppose the firm decides to issue more debt in the first stage (i.e. C_1 is higher). This means that equityholders will bear less of the investment cost (i.e. $F_1 - D_1$ is lower). This reduction in the equity-financing of the exercise cost enables equityholders to expedite the exercise of the first option. As a result, the firm produces first-stage cash flows, in expectation, earlier, which translates into a higher initial equity value, E_0. A higher C_1, however, reduces financial flexibility going forward. Therefore, the firm will have to wait longer to exercise the second option and to receive second-stage cash flows, which translates into a lower initial equity value, E_0. Taken together, the level of C_1 has two opposing effects in the multi-stage model, and the optimal level strikes a balance between the two effects.

On the other hand, suppose the firm decides to issue more debt in the second stage (i.e. C_2 is higher). This means that the firm uses more debt to fund the second-stage investment cost and hence equityholders bear less of the investment cost (i.e. $F_2 - D_{22}$ is lower). This expedites the exercise of the second option and the expected present value of cash flows from the second stage becomes higher, which increases initial equity value, E_0. However, the firm is more likely to default on its debt when there is more second-stage debt, which lowers the expected present value of cash flows from the second stage. In addition, the anticipation of the higher second-stage debt level lowers financial flexibility in the first stage. This leads to a delay of the first-stage investment, which decreases initial equity value, E_0. So, the optimal level of C_2 also trades off two opposing effects on initial equity value that follow from the dynamic interactions between financing and investment decisions in the multi-stage model.

In essence, dynamic financing-investment interactions in the "adolescent" and "mature" stages lead to an intertemporal effect. The firm trades off reaping exercise (i.e. cash flow) benefits from issuing debt in stage 1 against retaining financial flexibility for funding more of the investment cost with debt in stage 2. Note that this intertemporal effect does not depend on bankruptcy costs or tax benefit, which are constant over time and hence cannot cause timing differences. As we will see, the intertemporal effect is largely determined by the structure of the firm's investment process.

To quantify the intertemporal effect, we select economically plausible base case parameter values. The initial cash flow level, X_0, is set to \$5 and the risk-free interest rate equals $r = 6\%$. The growth rate of X is $\mu = 1\%$ and the volatility of X is $\sigma = 25\%$. The cost of bankruptcy cost is $\alpha = 30\%$ and the effective corporate tax rate is $\tau = 10\%$.[9] The scales of the sequential investment option are $\Pi_1 = 1$ and $\Pi_2 = 1$, and the investment costs are $F_1 = \$100$ and $F_2 = \$200$.

Using this base case environment, Figure 2.1 illustrates the intertemporal effect by mapping debt coupon pairs, C_1 and C_2, into initial equity value, E_0, on the basis of equation (2.16). The figure reveals that E_0 is convex C_1 and C_2 and, in particular, that an *interior optimum* is clearly present. Thus, the intuition behind the intertemporal effect discussed above leads indeed to an optimal pair of (C_1, C_2) that corresponds to the highest attainable point of initial equity value on the surface.

Table 2.2 reports, for the base case environment, optimal capital structure choices, investment thresholds, default thresholds, and market leverage ratios at the first and at the second investment point. Market leverage is defined as the ratio of market value of debt over the sum of market value of debt and market value of equity. Panel A tabulates the optimization results for the single-stage model, while Panel B shows the corresponding outputs for the multi-stage model. One of the key differences between the first column in Panel B and the first two columns in Panel A is the role played by financial flexibility.[10] That is, the underutilization of debt capacity, which we can gauge by the difference in leverage ratios between the single-stage model (42%) and the multi-stage model

[9] This tax rate is lower than those used in other studies. We do this on purpose, because we want to limit tax effects and emphasize the intertemporal effect due to the structure of the investment process. In addition, structural estimation of the multi-stage model in Section 2.4 provides fairly low point estimates of the effective corporate tax rate.

[10] Note that in Panel A the benchmark model's investment cost, F, equals either \$100 or \$200 to make it comparable to the first or the second stage of multi-stage model in Panel B. The same applies to the choice of investment scales.

(28% and 38%), shows that the firm in the multi-stage model has a strong incentive to retain financial flexibility in the first stage. In addition, the firm in the multi-stage model continues to have a lower target leverage ratio in the second stage. That is, underutilization persists and leverage ratios are lower in both stages of the multi-stage model relative to the ones of an otherwise identical firm in the single-stage model.

It is remarkable that leverage does not vary with characteristics of the investment option in the single-stage model (Panel A). Hence the intuition for the underutilization of debt in the multi-stage model (Panel B) is closely related to the one for the intertemporal effect. On the one hand, using more debt to finance the first investment lowers equityholders' contribution to the first investment cost, so equityholders will invest earlier (i.e. at a lower investment threshold). On the other hand, the pre-existing debt issued in the first stage creates default risk and reduces financial flexibility in the second stage, so equityholders will invest later (i.e. at a higher investment threshold). These countervailing effects lead to a variety of realistic outcomes, which depend crucially on the structure of the investment process (i.e. the relative size of the investment scales, Π_1 and Π_2). In particular, our calibrated model shows that firms underutilize debt in the first stage (i.e. when young) to maximize the overall value of their series of financing and investing options. Absent any other frictions or imperfections, dynamic optimization can generate path dependent and persistent leverage ratios in a standard trade-off model with a sequence of irreversible investment opportunities.

The findings in this section imply that structural models without dynamic financing-investment interactions (1) overestimate target leverage ratios, and (2) can be misleading in that they imply a fixed target leverage ratio that is largely taken to be exogenous to the investment process. The latter point also suggests that there is no meaningful measurement of partial adjustment to target leverage without recognition of the firm's investment process. Therefore, we will examine the role of heterogeneity in the structure of the investment process in the next section.

2.3.2 Structure of the Investment Process

To study how leverage changes with the structure of the investment process, we modify the base case of $\Pi_1 = 1$ and $\Pi_2 = 1$ (see first column in Panel B of Table 2.2) in columns 2–4. Column 2 depicts optimization results for a back-loaded investment structure ($\Pi_1 = 0.75$ and $\Pi_2 = 1.25$), while column 3 contains a front-loaded one ($\Pi_1 = 1.25$ and $\Pi_2 = 0.75$). Finally, column 4 reports the corresponding outcomes for a very front-loaded investment structure ($\Pi_1 = 1.5$ and $\Pi_2 = 0.5$).[11]

These columns highlight several interesting features of the model. First, the firm retains less (more) financial flexibility in the first stage when the structure of the investment process is front-loaded (back-loaded). For example, if we reduce the first-stage investment scale by 25% and raise the second-stage investment scale by 25%, then leverage in stage 1 decreases substantially from 28% to 20% (a decline of almost a third) and leverage in stage 2 increases from 38% to 40%. This result for the first stage helps explain the debt conservatism puzzle (see Graham (2000)).

Second, the difference in target leverage ratios across the two stages declines (rises) when the structure of the investment process is front-loaded (back-loaded). Consider the case in column 2, where 25% of the first-stage investment scale are pushed into the second stage, so the structure of the firm's investment process is more back-loaded than in the base case of column 1. As a result, the difference in target leverage ratios effectively doubles (i.e. increases from 10% to 20%). That

[11] Changing Π_1 and Π_2 in lock step produces the same coupons and leverage ratios as in column 1. All else equal, only the *ratio* of Π_1 and Π_2 matters because of of the scaling property. So we only consider asymmetric changes.

is, within our multi-stage model, firms that are otherwise identical (i.e. without considering the structure of investment process) need not follow the same target leverage ratios.

Third, the results in Panel B indicate that typically first-stage leverage is lower than second-stage leverage, which is consistent with dynamic capital structure models without investment (see, e.g., Goldstein, Ju, and Leland (2001) and Strebulaev (2007)). However, a sufficiently front-loaded investment process produces higher leverage in stage 1 than in stage 2 (see column 4). Intuitively, there is very little incentive to retain financial flexibility in this case and hence the firm utilizes its debt capacity more aggressively already in the first stage. Thus, unlike dynamic capital structure models without investment, the multi-stage model can produce increasing and decreasing leverage ratios over time, which are largely driven by the structure of the firm's investment process.

Finally and related to the previous point, it is perhaps also surprising that the mature firm with front-loaded investments selects lower leverage in the second stage (see columns 3 and 4) relative to the base case (see column 1). Another way to achieve lower leverage ratios in later stages would be to increase the firm's asset risk (e.g., because mature markets tend to be more competitive and hence riskier). It turns out, however, that even when the firm's asset risk is constant over time leverage ratios can nevertheless decline over time.

Overall, the analysis shows how incentives to retain financial flexibility in the first stage crucially depend on the structure of the investment process. The lower the first-stage investment scale, Π_1, is relative to Π_2, the more the firm saves debt capacity in the first stage. Given the wide range of optimal (target) leverage ratios, the results in Table 2.2 suggest that leverage ratios can greatly vary depending on how the firm grows assets-in-place by exercising its real options. Therefore, it is difficult to determine target leverage in the conventional sense unless the structure of investment process is recognized more explicitly in future empirical studies on target leverage and speed of adjustment to target leverage. In addition, corporate financing-investment dynamics in the multi-stage model produce a significant fraction of low (and zero) leverage firms and also path dependent, persistent leverage ratios (see also Section 2.6).

2.3.3 Other Comparative Statistics

We implement a sensitivity analysis to analyze how dynamic financing-investment interactions and leverage ratios change with parameter values unrelated to the investment process. We therefore vary the effective corporate tax rate, τ, the cost of bankruptcy, α, the cash flow volatility, σ, and the cash flow growth rate, μ. In particular, we increase and decrease each parameter by 50% of its base case value, while keeping the other parameters unchanged. Similar to Table 2.2, Table 2.3 reports optimal debt coupons, investment thresholds, default thresholds, and leverage ratios.

Not surprisingly, leverage goes up in both stages if the effective corporate tax rate, τ, rises. This follows from trade-off theory, because higher taxes lead to higher tax benefit of debt. The firm issues more debt and hence defaults earlier. The net effect of more debt issuance and more default risk is that optimal investment thresholds do not respond much to higher tax rates. The bankruptcy cost, α, has a negative impact on the value of debt, so it works in the opposite direction. Because higher bankruptcy costs lower the firm's debt capacity to fund the investment expenditure, investment takes place, in expectation, later. Yet, similar to the role of taxes, bankruptcy costs are also more related to debt and hence investment thresholds do not change significantly. Finally, notice that leverage ratios are sensitive to both changes in bankruptcy costs and changes in tax rates.

Since σ is a measure of uncertainty, real option theory tells us that the investment option's value rises with σ and that, as a result, exercise should occur, in expectation, later. In the multi-stage model, initial equity value equals the value of a sequential investment (compound) option. Therefore, changes in volatility greatly change investment decisions. When volatility goes up, it is more valuable to keep the option alive, so the firm waits longer to invest in both stages (X_{S1} rises from 12.36 to 17.51 in stage 1 and X_{S2} rises from 23.67 to 33.08 in stage 2 for the case in column 6). Consistent with many studies building on Leland (1994), varying volatility also leads to significant variation in leverage. When volatility is high, the firm is riskier (i.e. more likely to go bankrupt) and hence uses debt more conservatively in both stages and leverage ratios drop significantly relative to the base case. Thus, it can pose a challenge for structural estimation that volatility strongly affects both financing and investment decisions.

Finally, a higher growth rate of cash flows, μ, also makes the firm's options more valuable, because their intrinsic value is higher for any level of cash flows and hence waiting to invest becomes costlier. A rise in μ therefore leads to, in expectation, earlier investment (i.e. X_{S1} declines from 12.36 to 11.99 in stage 1 and X_{S2} declines from 23.67 to 22.91 in stage 2 as seen in column 8). A rise in μ also leads the firm to use more debt to finance investment in both stages. However, leverage ratios are fairly insensitive to 50% increases or decreases in the growth rate, which might also make structural estimation of μ more challenging.

In sum, the multi-stage can, for reasonable base parameter values, deliver leverage ratios and variation in leverage ratios, which closely match those observed in practice. In the next section, we tackle the more demanding question whether several simulated model moments can simultaneously match a number of desirable data moments. While the sensitivity analysis sheds light on multi-stage model, it also helps us find informative moments for different model parameters in the next section.

2.4 Simulated Moments Parameter Estimation

We now turn to simulation methods based on indirect inference techniques in Gourieroux, Monfort, and Renault (1993) and Gourieroux and Monfort (1996). Like Hennessy and Whited (2005, 2007), we use Simulated Method of Moments (SMM) to estimate a set of structural parameters of the model. To do so, we solve the multi-stage model numerically and then use this solution to generate simulated panels of firms. SMM selects parameter values by minimizing the distance between moments from actual data and corresponding moments from simulated data. That is, the goal of SMM is to find an optimal vector of unknown structural parameters, say b^*, by matching a set of *simulated moments*, denoted by M_m, with corresponding *data moments*, denoted by M_d.

Let $b = (\Pi_1, \Pi_2, \alpha, \tau)$ be the vector of unknown structural parameters to be calibrated by SMM. We simulate $S = 6$ artificial panels data sets consisting of $N = 1,000$ firms for 200 years using the multi-stage model and a given b. In each panel, we only keep $T = 100$ years (or 400 quarters given that $\Delta t = 0.25$) after discarding the first 100 years to avoid undue influence of initial conditions. Once a firm defaults, we replace it by a new firm with the same characteristics to keep the size of the simulated data sets constant over time. By iterating b and calculating the distance between the model moments, M_m, and data moments, M_d, SMM returns the optimal vector of parameter values, b^*. We repeat this indirect inference procedure ten times and report the average of the ten SMM results for b^* to reduce noise.[12]

The moments to match are selected such that they are *a priori* informative about the unknown structural parameters, b. Intuitively, a moment is informative about b if it can identify at least some elements in b, which means it is sensitive to changes in b. Informative moments enable SMM to

[12]See Hennessy and Whited (2005, 2007) on the benefits and relevance of indirect inference techniques in finance.

converge faster and to provide more robust economic insights. To this end, we select the following five moments to estimate the four structural parameters $(\Pi_1, \Pi_2, \alpha, \tau)$:[13]

1. Quasi-Market Leverage (QML): Let QML_t^i denote the quasi-market leverage ratio (i.e. the book value of debt divided by the sum of market value of equity and book value of debt) of the simulated firm i, with $i = 1...N$, at time t, with $t = 1..., T$. We first calculate the cross-sectional average for every time t and then take the average of the cross-sectional averages over time, i.e. $QML = \frac{1}{T}\sum_{t=1}^{T}(\frac{1}{N}\sum_{i=1}^{N} QML_t^i)$. This moment reflects how the net benefits of debt and, in particular, the structure of the investment process affects the level leverage. So we expect it to be sensitive to, e.g., Π_1 and also responsive to α and τ.
2. Dispersion of Quasi-Market Leverage ($DispQML$): This moment is defined as the cross-sectional average of the time-series standard deviations of firms' quasi-market leverage ratios, i.e. $DispQML = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\text{Var}_i(QML_1^i, ..., QML_T^i)}$. This moment reflects how firms respond to shocks when optimally financing their investment projects. Hence this moment is also likely to be informative about Π_1 as well as α and τ, but it captures cross-firm variation in leverage.
3. Debt Issuance (D/K): We compute D/K as the ratio of net debt issuance to capital at investment points. This moment reflects the proportion of debt used to finance investment expenditures, which is higher if bankruptcy costs are lower or if tax rates are higher. It should be sensitive to the parameters α and τ that determine the net benefits of debt issuance.
4. Market-to-Book (Q): We first calculate the cross-sectional average for every time t and then take the average of the cross-sectional averages over time, i.e. $Q = \frac{1}{T}\sum_{t=1}^{T}(\frac{1}{N}\sum_{i=1}^{N} Q_t^i)$, where Q_t^i is the market value of firm i divided by the book value of firm i at time t. Q proxies for investment opportunities and hence it should be informative about Π_1 and especially Π_2.
5. Investment-to-Equity (Inv/Eq): This moment is defined as investment cost (i.e. F_1 or F_2) divided by book value of equity at investment points. Since the investment costs are fixed, this moment depends more on when the firm exercises its options and, in particular, how high equity value has to rise for exercise to be optimal. Therefore, it is likely related to Π_1 and Π_2, but potentially also to τ because equity is a claim on after-tax cash flows.

Table 2.4 presents the sensitivities of the five moments to changes in the model parameters b. The base case scenario is in the first column. In the other columns, each of the four parameters, namely Π_1, Π_2, α and τ, is separately increased by 50%. Panel A displays the absolute changes of the moments and Panel B shows percentage changes relative to the base case in the first column. The table reveals that QML and $DispQML$ are indeed most sensitive to Π_1. This is because QML reflects the investment and financing activities in the past, i.e. the level and dispersion of early stage leverage ratios are primarily determined by how high Π_1 is. Tax rate τ and bankruptcy cost α most strongly affect the debt issuance moment, D/K, because they are the key determinants of both debt capacity and net tax benefits. Q is very sensitive with respect to changes in Π_2 Recall that initial equity value corresponds to a compound option in the multi-stage model. All else equal, firms with more a back-loaded investment process (i.e. higher Π_2) will have much higher market-to-book ratios. The investment to book equity ratio Inv/Eq responds the most to τ and Π_2 and also a bit less so to Π_1. In sum, Π_1 influences mainly QML and $DispQML$, Π_2 matters the most for Q and also changes D/K, α and τ have the strongest effect on D/K but also affect QML, and finally τ (but not α) impacts Inv/Eq.

We use the COMPUSTAT annual tape for the 45-year period between 1965 and 2009 to estimate the data moments.[14] We refer to this vector as the "Full Sample" moments. We run SMM also on

[13] Appendix B provides a more detailed description of the structural estimation procedure that we implement here.

[14] We remove financial firms (SIC between 6000 and 6999) and utilities (SIC between 4900 and 4999), because they operate in regulated industries. We also remove firm-year observations with total assets less than two million and plant, property, and equipment less than one million to avoid biases caused by small firms.

various subsamples to provide insights into the structure of companies' investment process. Therefore, the COMPUSTAT sample is split into three subsamples according to Q. For each firm-year, firms are ranked by the value of Q from the lowest to the highest. Firms in the lowest 33% of the distribution are classified as the "Low Q" sample, firms with the highest 33% Q belong to the "High Q" sample, and those in between are the "Medium Q" sample. For each of the subsamples, we also compute the corresponding data moments. We then run SMM for each of the four samples and collect the vector of structural parameter estimates, b^*, along with the (fitted) model moments.

The parameter estimates for the full sample and the three subsamples are presented in Panel A of Table 2.5. The numbers in the parentheses are the standard deviations of model estimates across the 1,000 panels with the exception of the χ^2 test, for which that number is the p-value for the overidentification test. In particular, the χ^2 test of the overidentifying restrictions does not produce a rejection at the 10% level for any of the four structural estimations in Panel A. Panel B of Table 2.5 reports the data moments and the fitted (model) moments for the full sample and the three subsamples. For all four estimations, the fitted moments are very close to the data moments. This implies that the simulated economy with these parameter estimates closely mimics the real economy.

As argued earlier, the ratio of investment scales, Π_2/Π_1, captures the structure of investment process. Panel A reveals that the "High Q" sample has a fairly back-loaded investment processes ($\Pi_2/\Pi_1 = 2.284 > 1$), whereas the other three samples display various front-loaded investment processes. For example, the "Low Q" sample shows the most front-loaded investment process ($\Pi_2/\Pi_1 = 0.175 < 1$), which is also the only case of Π_1 not being reliably measured. Importantly, these results of the structural estimation are consistent with our theoretical discussion in Section 2.3.2.

Specifically, firms in the "High Q" sample tend to have many future investment opportunities and thus large growth potential — the structure of the investment process of these firms is indeed back-loaded, so they initially retain more financial flexibility (i.e. save more debt capacity). This is why we observe the lowest leverage ratio for this sample. The "Low Q" sample has more mature firms and they do not have many future investment opportunities. Their investment process tends to be more front-loaded, so they do not have a motive to save debt capacity for future investments and this is why we observe the highest leverage ratio for this sample. On average, our 45-year COMPUSTAT sample contains more firms having fairly front-loaded investment process.

Another interesting observation is that the estimated tax rate, τ, is low for all four samples, ranging from 4% to 7%. There is more variation in the estimates of the bankruptcy costs, α, but if we put less weight on the "Low Q" sample, then it is also roughly 30%. This implies that the net tax benefit of debt is not driving our result, confirming the conclusion of Section 2.3. While, these findings are also consistent with the debt conservatism puzzle in that the data moments of observed leverage ratios are in line with relatively low effective tax rates, they establish, more importantly, that the structure of the investment process is likely to be a more important determinant of leverage ratios than tax rates or bankruptcy costs, because the SMM indicates that its parameters, Π_1 and Π_2, vary much more widely across subsamples than α and τ.

In sum, unobserved heterogeneity in terms of the precise structure of the investment process (e.g. front-load, mid-loaded, or back-loaded) seems to be large and important for understanding the cross-sectional distribution of corporate investment and financing decisions within an industry or for the entire economy. Thus, future capital structure research should try to move more into the direction of recognizing and treating this important source of cross-firm heterogeneity.

2.5 Capital Structure Regressions

Simulation is a useful tool for testing theoretical models (see, e.g., Hennessy and Whited (2005), Leary and Roberts (2005), Strebulaev (2007), and Tserlukevich (2008)). In this section, we investigate the cross-sectional properties of leverage ratios in dynamic, simulated economies where firms make infrequent investment and financing decisions. We also compare capital structure regression results for simulated data to the corresponding results for real data.

2.5.1 Panel Simulations

This section outlines the details of simulating dynamic economies with heterogeneous firms. These simulations take the solutions for the valuation equations and, in particular, for the optimal investment and financing decisions in Section 2.2.2 as given and do not involve any additional optimizations.

It is well-known that systematic (or economy-wide) and idiosyncratic (firm-specific) shocks have explanatory power for leverage ratios. The former create cross-firm dependencies in dynamic economies. Hence we decompose the cash flow process into systematic and idiosyncratic components to allow for a more realistic correlation structure:

$$dX_i(t)/X_i(t) \;=\; \mu\, dt+\beta_i\, \sigma_S\, dW^S(t)+\sigma_{Ii}\, dW_i^I(t),\; X_i(0) = X_0 > 0,\; i = 1, ..., N,\; t = 1, ..., T \quad (2.24)$$

where constants σ_S and σ_{Ii} represent, respectively, volatilities of systematic and idiosyncratic shocks. The stochastic processes $dW^S(t)$ and $dW_i^I(t)$ are independent Wiener processes, $dW(t) = \beta_i\, dW^S(t) + dW_i^I(t)$, and the parameter β_i measures firm i's exposure to systematic shocks. Thus, the total risk can be calculated as $\sigma_i = (\sigma_{Ii}^2 + \beta_i^2\sigma_S^2)^{\frac{1}{2}}$.

Based on a discretization of equation (2.24), we simulate 1,000 panel data sets with $N = 3,000$ firms for $T = 400$ quarters. More specifically, we generate 200 years of data (or 800 quarters given that $\Delta t = 0.25$) based on the multi-stage model and then drop the first 400 quarters to obtain a stationary sample for each simulated economy and to limit the influence of initial conditions. All firms in the same panel are governed by the same series of systematic shocks, $dW^S(t)$, but the dynamics of economies are different across the 1,000 panels. At time $t = 0$ of each panel, the optimal policies are determined as a function of firm characteristics (i.e. parameter values). For all $t > 0$, firms follow their optimal investment, debt and equity issuance, and default policies given that they observe the evolution of their cash flow process every quarter. If a firm defaults, it is "reborn," i.e. it is replaced in the next quarter by a juvenile firm with identical initial conditions.

Panel A of Table 2.6 provides an overview of the values and distributions of model parameters that are set once and for all at time $t = 0$ to produce the simulated economies. To begin, we use the structural estimation results for the full sample in Section 2.4. That is, the investment scales are calibrated for all firms to $\Pi_1 = 1.966$ and $\Pi_2 = 1.286$. The bankruptcy cost, α, and the tax rate, τ, are uniformly distributed with means corresponding to the SMM's full sample results and ranges of +/–20% around the means. Similarly, the investment costs, F_1, and F_2, have uniform distributions with supports $[80, 120]$ and $[160, 240]$, respectively. The systematic shock is fixed to $\sigma_S = 0.148$ based on estimates reported by Schaefer and Strebulaev (2008). The idiosyncratic shock, σ_{Ii}, varies around a mean of 0.217 based on a chi-squared distribution: $\sigma_{Ii} \sim 0.05 + \frac{1}{30}\, \chi^2(5)$. Firm i's exposure to systematic shocks, β_i, follows a uniform distribution, whose first two moments correspond to the empirical distribution with mean of 0.993 and standard deviation of 0.47 reported by Strebulaev

(2007). Finally, the other parameters assume the base case values, i.e. risk-free rate $r = 6\%$, growth rate of cash flows $\mu = 1\%$, and initial cash flow level $X_0 = \$5$.

Panel B of Table 2.6 presents the cross-sectional distribution of leverage both at investment points and across all panels (i.e. in dynamics). For each simulated data set, we first calculate statistics for each quarter. We then average across quarter within each simulated economy, and then average across economies. However, the rows "Min." and "Max." report, respectively, the minimum and the maximum assumed by the corresponding quantities. Investment points are further classified as first and second investment points, because the multi-stage model features two investment options. Market leverage, ML, is defined in Section 2.3.1. Quasi-market leverage, QML, equals book value of debt divided by the sum of market value of equity and book value of debt in this model, where book value of debt is defined as the value of debt at the beginning of a stage (i.e. at $X = X_{S1}$ or $X = X_{S2}$). Hence quasi-market leverage and market leverage coincide at investment points.

The table confirms that ML and QML are not very different. Across all panels, the average ML is 19.7% and the average QML is only 0.8% higher. Yet, the market leverage ratios at first investment points are almost half of those at second investment points, which is attributable to the strong incentive of adolescent firms to retain financial flexibility. Taken together, average leverage at investment points equals 18.7%. This is significantly lower than in similar models without endogenous investment by Goldstein, Ju, and Leland (2001) and Strebulaev (2007), who report 37% and 26%. Hackbarth and Mauer (2012) obtain optimal leverage ratios as low as 12%, but their success is largely due to debt overhang and debt (dilution) dynamics, whereas our simulations do no rely at all on agency problems. Moreover, the standard deviations of leverage ratios at investment points are about half of those for all data points. This is due to the fact that firms tend to make infrequent investments that occur at optimal times. Thus, even though target leverage ratios differ across firms, they are less dispersed than leverage ratios in dynamics. In addition, the distribution of leverage ratios at investment points are almost symmetric, whereas the ones in dynamics are right-skewed (i.e. the mean exceeds the median). Intuitively, because firms in various (investment) stages of the model or, more generally, their life-cycle, respond differently to economic shocks of the same magnitude, the behavior of leverage ratios at investment points is quite different from that in dynamics.

In sum, the simulations of heterogeneous firms generate much lower average leverage ratios than prior work, both at the investment points and in dynamics, and yet it is still able to generate leverage ratios spanning over the $[0, 1]$ interval. Therefore, we conclude that the structure of firms' investment process is crucial for obtaining realistic distributions of leverage ratios in simulated economies. Intuitively, firms endogenously preserve debt capacity (i.e. retain financial flexibility) for future investment stages. That is, the intertemporal effect of the multi-stage model of corporate investment and financing dynamics captures an important mechanism which helps explain the low-leverage puzzle of Graham (2000). More recently, Strebulaev and Yang (2012) document a closely related, so-called zero-leverage puzzle given that, e.g., 14.0% of large, public companies had no debt outstanding in the year 2000. Consistent with their findings, our model produces, on average, also produces a large and persistent fraction of zero-leverage firms. As seen from the 10th percentile in Panel B of Table 2.6, the fraction of zero-leverage firms in dynamics exceeds, on average, 10% for the 1,000 simulated panel data sets. Clearly, this suggest that, at the very least, the model is able to explain a substantial part, if not most, of the low- and zero-leverage puzzles. It seems that this success cannot be achieved by alternative models without endogenous investment, so considering dynamic interactions between corporate investment and financing is crucial.

2.5.2 Panel Regressions

In this section, we focus on the behavior of quasi-market leverage, market-to-book, and profitability in simulated data sets. In particular, we estimate capital structure regressions both at investment points and for all observations in the panels (i.e. in dynamics). This also allows us to examine the role of other determinants of capital structure that are typically used in the empirical literature.

Growth options might have a negative debt capacity, because debt overhang rises with leverage (Myers (1977)). Indeed, numerous empirical studies find a negative relation between leverage and market-to-book, a commonly used proxy for growth options, and interpret it as evidence for agency problems. For example, Smith and Watts (1992) document a negative relation between quasi-market leverage ratios and market-to-book ratios. Similarly, Rajan and Zingales (1995) report a reliably negative relation between quasi-market leverage and the market-to-book ratio across seven different countries. However, they conclude: "From a theoretical standpoint, this evidence is puzzling. If the market-to-book ratio proxies for the underinvestment costs associated with high leverage, then firms with high market-to-book ratios should have low debt..." Yet, Chen and Zhao (2006) find leverage is positively related to market-to-book for all firms except those with the highest market-to-book ratios. We therefore examine the leverage-growth relation in our simulated economies.

A distinct yet related line of research studies the relation between leverage and profitability. Myers (1993) argues that the negative relation between leverage and profitability is one of the most pervasive patterns of empirical capital structure research. According to Strebulaev (2007), it is also one of a few relations that enables us to distinguish between the (static and dynamic) trade-off model and pecking order behavior. We therefore examine whether conclusions from prior research on the leverage-profitability relation in dynamic capital structure models without investment carry over to our simulated economies, in which corporate investment and financing are endogenous.

More specifically, the empirical variables of interest are profitability, π, and market-to-book, Q. The interaction of leverage with these two factors is widely used to differentiate implications of the trade-off theory of capital structure from the pecking order. Based on the standard trade-off theory, higher profitability enables firms to reduce the costs of bankruptcy and increase the tax benefit of debt. Thus, a positive leverage-profitability relation is predicted. This prediction is challenged by a large body of empirical research, such as Titman and Wessels (1988), Rajan and Zingales (1995), and Fama and French (2002), and Baker and Wurgler (2002), who all find a negative association confirming the pecking order's prediction. Firm behavior according to the pecking order means that higher profitability allows firms to use more retained earnings. Holding investment fixed, leverage is thus lower for more profitable firms. As a result, the negative leverage-profitability relation has traditionally been regarded as evidence in favor of pecking-order and against trade-off behavior. Regarding the leverage-growth relation, these two theories have opposite predictions too. In the trade-off world, high growth firms tend to have lower collateral values and hence higher bankruptcy costs. Trade-off firms with high growth should therefore issue less debt. In the pecking-order world, debt increases when capital expenditures are higher than retained earnings, and decreases when capital expenditures are lower than retained earnings. Holding profitability fixed, leverage is thus higher for pecking-order firms with high growth. Taken together, a positive (negative) leverage-Q (or leverage-π) relation follows from pecking order (trade-off) arguments.

Recall that we generate 1,000 panel data sets with 800 quarters based on the multi-stage model and then drop the first 400 quarters to obtain a stationary sample for each simulated economy and to limit the influence of initial conditions. Using these simulated economies, we estimate four versions of a standard capital structure regression model for quasi-market leverage:

$$QML_i = \beta_0 + \beta_1\, x_i + \beta_2\, \sigma_i + \beta_3\, \alpha_i + \beta_4\, \tau_i + \epsilon_i, \tag{2.25}$$

where x is either profitability, π, in Panel A of Table 2.7 or market-to-book, Q, in Panel B of Table 2.7. In Panel C of Table 2.7, we include both profitability, π, and market-to-book, Q, as regressors. We measure profitability, π, as earnings before interest and tax (or cash flows) scaled by firm value, whereas the market-to-book ratio, Q, is the ratio of total market value of asset over book value of asset. The other independent variables are time-invariant firm attributes: volatility of cash flows, σ, bankruptcy cost, α, tax rate, τ, and firm size, φ, which equals the sum of book value of debt and book value of equity. We focus on QML as regressand, because distributions of market leverage and quasi-market leverage closely mimic each other in our simulated economies (see Table 2.6).

In Table 2.7, the first column reports the regression results at investment points and the other columns report the ones on simulated economies.[15] In particular, the first version of equation (2.25) is the "Investment Points" regression, whose estimation results are tabulated in the first column (Inv. Pts.) of Table 2.7.[16] The second column (BJK) of Table 2.7 reports OLS regression results in the fashion of Bradley, Jarrel, and Kim (1984). The dependent variable, QML_i, is calculated as the sum of book values of debt over the 400 quarters divided by the sum of book values of debt and market values of equity over the 400 quarters. The independent variables are calculated similarly (if possible). Note that for this regression the dependent variable and independent variables are contemporaneous. In the third column (RZ), the independent variables are averaged over quarters $t-1$ to $t-4$ to reduce noise as in Rajan and Zingales (1995), while the dependent variable, QML_i, is measured at time t. Finally, the fourth column (FF) adopts the Fama-MacBeth regression approach as in Fama and French (2002). At each time t, QML is regressed on lagged independent variables. We report averages of the quarter-by-quarter coefficient estimates along with Fama-MacBeth standard errors that are corrected with the Newey-West method.

Strebulaev (2007) points out empirical capital structure regressions should differentiate refinancing points from other data points. He develops a dynamic trade-off model with financing frictions, where firms size if fixed over time. An important question is therefore whether similar conclusions obtain if firm size is not fixed over time, so firms can make a sequence of optimal investment and financing decisions. To this end, Table 2.7's Panel A reveals that the leverage-profitability relation is positive and significant at investment points (see first column), which is consistent with static trade-off behavior. Interestingly, it is reliably negative in the other columns, in which we estimate BJK, RZ, and FF regressions using all data points (i.e. in dynamics). This is the pecking order's prediction, but derives entirely from data produced by a dynamic trade-off model with investment frictions. The effects of cash flow volatility, bankruptcy cost, tax rate, and firm size on leverage are also significant and go in the expected directions. The upshot of Panel A is that infrequent, lumpy investment provides an economically important alternative to financing frictions, because we observe remarkably different patterns at investment points versus in dynamics.

Furthermore, the intertemporal effect of the multi-stage model has implications for the leverage-growth relation, which we examine in Panel B of Table 2.7. For example, Frank and Goyal (2009) summarize that market-to-book has a reliably negative relation with leverage, which is consistent with the prediction of both trade-off theory and market-timing theory. Absent agency problems or

[15] Coefficient estimates and t-statistics reported in this table are the averages across the 1,000 simulated economies.

[16] In Strebulaev (2007), the first regression is run at refinancing points only as he considers only financing friction. Our paper considers only investment frictions. In reality, one would, however, expect both frictions to be important.

market inefficiencies, this phenomenon is strongly borne out in the regressions on all data points (i.e. in dynamics). Again, the sign is reversed at investment points. Thus, the interpretation of cross-sectional tests of the leverage-Q relation changes depending on whether firms are active (i.e. at investment points) or passive (i.e. in between investment points). While firm size, φ, and firm risk, σ, are both economically and statistically significant in Panel B, bankruptcy cost, α, and tax rate τ are less reliable variables, consistent with Frank and Goyal (2009). Consider, for example, the RZ regressions, where α and τ are not even 10% significant. In contrast, α and τ are highly significant in Strebulaev (2007), where Q is not included as independent variable in the regression analysis.

Finally, Panel C includes both profitability, π, and market-to-book, Q, as regressors. The results are similar to the ones in Panels A and B. Interestingly, neither π nor Q lose statistical power, even though both are influenced by the same underlying sources of uncertainty. Thus, this last part of our regression analysis suggests that profitability and market-to-book are independently important for explaining the behavior of leverage ratios. More generally, we expect in reality both financing frictions and investment frictions. So, these complementary types frictions will be present at different points in time (i.e. at separate investment and refinancing points) and also at the same points in time as in our model. Clearly, this can only strengthen the relevance of our conclusions.

To summarize, the regression results based on model simulated data using our structural estimation results are able to replicate stylized facts established by empirical research. In the spirit of Strebulaev (2007), simulation of the multi-stage model of corporate investment and financing dynamics reinforces the need to differentiate investment points from other data points when studying corporate behavior in a dynamic world.

2.6 Capital Structure Persistence

In a recent article, Lemmon, Roberts and Zender (2008, LRZ) chart the future evolution of leverage ratios for four portfolios constructed by sorting firms based on their actual leverage (Figure 1) or unexpected leverage (Figure 2). LRZ report the puzzling evidence that, despite of some convergence, average leverage ratios across the four portfolios are fairly stable over time for both types of sorts (i.e. actual and unexpected leverage). During the 39 year period from 1965 to 2003, firms with relatively high (low) leverage tend to maintain relatively high (low) leverage for over 20 years. LRZ conclude that the striking stability in leverage paths is unexplained by previously identified determinants (e.g., firm size, profitability, market-to-book, industry, etc.) or changes in sample composition (e.g., firm exit).

In this section, we establish that investment frictions in a dynamic model can produce average leverage paths that closely resemble the corresponding ones documented by LRZ. Put differently, optimal investment and financing decisions in a dynamic model can largely explain the otherwise puzzling patterns. To see this, we generate simulated data for an extension of our multi-stage model. At time $t = 0$, we introduce an initial coupon, C_0, that will create variation in initial leverage. But all firms have the same initial scale, Π_0, so they can service this initial debt issue. Given that each firm has debt in place in stage 0, there is also an endogenous default threshold, X_{D0}. The values of initial debt in stages 0, 1, 2 are, respectively, denoted by D_0, D_{10}, and D_{20}. All other variables are the same as in Section 2.2 (see Appendix A for more details).

We proceed again by generating simulated economies. In particular, the sample consists of 3,000 firms over 39 years in 1,000 panel data sets with three industries, which have 1,000 firms each and are defined based on the subsample (i.e. low, medium, high Q) estimation results for b^* in Section

2.4. The modeling of firm-level heterogeneity follows the procedure in Section 2.5.1, except that we replace the SMM's full sample by the SMM's three subsample estimation results. For example, the "Medium Q" industry has investment scales for all firms equal to $\Pi_1 = 2.032$ and $\Pi_2 = 0.848$, the bankruptcy cost, α, and the tax rate, τ, are centered around 0.267 and 0.047, respectively, based on a uniform distribution with ranges of +/–20% around their centers, etc. The initial scale is normalized to one, $\Pi_0 = 1$, but firms have an exogenously assigned initial coupon, C_0, which is drawn from a lognormal distribution: $C_0 \sim \text{LogNormal}(0.5, 1)$. For each point in time or quarter, $t = 0, ..., 156$, in the simulated economies, we compute book leverage and quasi-market leverage for each firm.[17]

For the simulated data, we implement the same procedure as outlined in LRZ to track average leverage ratios of firms in four portfolios, denoted by "Very Low", "Low", "High" and "Very High," based on these firms' actual leverage. Figure 2.2 presents the average book and quasi-market leverage ratios for these portfolios in "event time," both for all simulated firms in Panels A and B and for survivors (i.e. firms that exist at least for 20 years) in Panels C and D. In our simulated data, firms leave the sample because of bankruptcy. From quarter 76 onward, the length of time for which we can follow each portfolio is censored because we only simulate data for 156 quarters. Therefore, we follow LRZ and only perform the portfolio formation through quarter 76 in case of survivors.[18]

The figure shows for our model simulated data sets that average leverage paths for the four portfolios formed as in LRZ converge to stable levels in the long run. However, they do not converge to target leverage, which would be predicted by static trade-off models in which firms always converge to target as soon as they make adjustments (or as soon as adjustment costs allow them to do so). Recall that the analysis in Section 2.3 shows that firms that are otherwise identical (i.e. without considering the structure of investment process) need not follow the same target leverage ratios. That is, within our multi-stage model, firms that are otherwise identical (i.e. without considering the structure of investment process) need not follow the same target leverage ratios. Given that the structure of investment process is hard to observe perfectly, the persistence of leverage in Figure 2.2 therefore means that this unobservable heterogeneity can give the appearance of a transitory or short-term component of debt, even though firms dynamically optimize their permanent or long-term component of debt in our model by trading off net tax benefits, financial flexibility, and investment benefits.

LRZ note that a potential concern regarding their main finding is that constructing portfolios based on actual leverage can implicitly pick up cross-sectional variation in underlying factors, which themselves influence cross-sectional variation in leverage, such as bankruptcy costs or industry attributes. Like LRZ, we therefore also form four portfolios by ranking firms based on their "unexpected leverage" and then track again the portfolios' averages of actual leverage in event time. Unexpected leverage is defined as the residuals from a cross-sectional regression of leverage on market-to-book, Q, profitability, π, volatility of cash flows, σ, bankruptcy cost, α, tax rate, τ, firm size, φ, and industry indicator variables, where all independent variables are lagged one year.

Figure 2.3 presents the graphs for unexpected leverage portfolios. It shows that the results are nearly identical to those for actual leverage portfolios in Figure 2.2. While there is slightly less cross-sectional variation in average leverage, the differences in average leverage across the four portfolios do not quickly disappear in our simulated economies. Thus, unexpected leverage portfolios cannot remove the key variation in C_0 that creates the initial cross-firm differences and then as a result of large enough, real frictions the striking stability in average leverage paths for very long periods of time. So we conclude that persistence is not a special case of some parametrization, simulation, or

[17] We find qualitatively identical results if we simulate 139 years and drop the first 100 years or when C_0 obeys a uniform distribution: $C_0 \sim \text{Uniform}[0.01, 6]$. These unreported results are available from the authors on request.

[18] We suppress 95% confidence intervals in the figures, because they almost coincide with the average leverage lines.

sorting procedure but rather a general result of a dynamic trade-off model with investment frictions.

In sum, this section implies that the corporate investment process can be the driving force behind leverage ratios' pronounced persistence and remaining variation after long periods of time elapse. We thus need to focus more on the heterogeneity in companies' investment processes.

2.7 Conclusion

This paper examines interactions of corporate investment and financing in a dynamic trade-off model with a sequence of irreversible investment opportunities. The model produces an intertemporal effect, which links financial flexibility to the investment process. Firms trade off reaping investment (i.e. cash flow) benefits from issuing debt in an earlier stage against retaining financial flexibility for funding more of the investment cost with debt in a later stage. Taking future financing and investment opportunities into account, (juvenile) firms tend to underutilize debt when financing investment the first time to retain financial flexibility. Surprisingly, the underutilization of debt persists even when the adolescent firm matures (i.e. exercises its last investment opportunity). In addition, underutilization of debt is more (less) severe for the more back-loaded (front-loaded) investment opportunities. That is, even within a standard trade-off model, firms that are identical without considering their investment opportunities do not follow the same target leverage ratios.

Parameter estimation via Simulated Method of Moments takes the model to the real data. Structural estimation results reveal that high growth firms have, on average, a more back-loaded investment process, which helps explain why they tend to have low leverage ratios. Furthermore, capital structure regression results for model simulated data using these estimation results produce stylized facts consistent with the empirical literature. Notably, our dynamic trade-off model with a sequence of irreversible investments is capable of producing a negative leverage-profitability relation and, in the absence of agency problems or other frictions, a negative leverage-growth relation. Therefore, empirical tests without incorporation of the structure of the investment process (and in particular cross-firm variation thereof) can be uninformative to the extent that their interpretation is not robust to recognizing heterogeneity in companies' investment opportunities. Finally, an extension of our dynamic framework to randomly imposed initial variation in leverage reveals that the model can fully explain the empirical puzzle that average observed leverage ratios are path dependent and persistent for very long periods of time.

Overall, we conclude that it is important for studies of capital structure to recognize the structure of the investment process, which strongly influences both investment and financing dynamics. The rather rich set of insights and predictions generated by embedding a sequence of irreversible investments in a dynamic trade-off model suggests that further extension of this class of models will prove fruitful for future research. prove fruitful for future research.

2.8 References

Adam, Tim, and Vidhan K. Goyal, 2008, Investment opportunity set and its proxy variables, *Journal of Financial Research* 31, 41–63

Baker, Malcolm, and Jeffrey Wurgler, 2002, Market timing and capital structure, *Journal of Finance* 57, 1–32.

Barclay, Michael J., Morellec, Erwan, and Clifford W. Smith Jr., 2006, On the debt capacity of growth options, *Journal of Business* 79, 37–59.

Bradley, Michael, Jarrell, Gregg A., and E. Han Kim, 1984, On the existence of an optimal capital structure: Theory and evidence, *Journal of Finance* 39, 857–878.

Chen, Long, and Xinlei Zhao, 2006, On the relation between the market-to-book ratio, growth opportunities, and leverage ratio, *Finance Research Letters* 3, 253–266.

DeAngelo, Harry, Linda DeAngelo, and Toni M. Whited, 2009, Capital structure dynamics and transitory debt, Working Paper, University of Rochester.

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

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

Fischer, Edwin O., Robert Heinkel, and Josef Zechner, 1989, Optimal dynamic capital structure choice: Theory and tests, *Journal of Finance* 44, 19–40.

Flannery, Mark J., and Kasturi P. Rangan, 2006, Partial adjustment toward target capital structures, *Journal of Financial Economics* 79, 469–506

Frank, Murray, and Vidhan K. Goyal, 2003, Testing the pecking order theory of capital structure. *Journal of Financial Economics* 67, 217–248.

Frank, Murray Z., Vidhan K. Goyal, 2009, Capital structure decisions: Which factors are reliably important?, *Financial Management* 38, 1–37

Goldstein, Robert, Ju, Nengjiu, and Hayne E. Leland, 2001, An EBIT-based model of dynamic capital structure, *Journal of Business* 74, 483–512.

Gourieroux, Christian, and Alain Monfort, 1996, *Simulation Based Econometric Methods* (Oxford University Press, Oxford, U.K.).

Gourieroux, Christian, Monfort, Alain, and Eric Renault, 1993, Indirect inference, *Journal of Applied Econometrics* 8, S85–S118.

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

Hackbarth, Dirk, Hennessy, Christopher A., and Hayne E. Leland, 2007, Can the trade-off theory explain debt structure?, *Review of Financial Studies* 20, 1389–1428.

Hackbarth, Dirk, and David C. Mauer, 2012, Optimal priority structure, capital structure, and investment, *Review of Financial Studies* 25, 748–796.

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

Leary, Mark T., and Michael R. Roberts, 2005, Do firms rebalance their capital structures?, *Journal of Finance* 60, 2575–2619.

Leland, Hayne E., 1994, Corporate debt value, bond covenants, and optimal capital structure, *Journal of Finance* 49, 1213–1252.

Lemmon, Michael L., Roberts, Michael R., and Jaime F. Zender, 2010, Back to the beginning: Persistence and the cross-section of capital structure, *Journal of Finance* 63, 1575–1608.

Modigliani, Franco, and Merton Miller, 1958, The cost of capital, corporation finance and the theory of investment, *American Economic Review* 48, 261–297.

Morellec, Erwan, and Norman Schuerhoff, 2011, Corporate investment and financing under asymmetric information, *Journal of Financial Economics* 99, 262-288.

Myers, Stewart C., 1977, Determinants of corporate borrowing, *Financial Economics* 5, 147–175.

Myers, Stewart C., 1993, Still searching for optimal capital structure, *Journal of Applied Corporate Finance* 6, 4–14.

Rajan, Raghuram G., and Luigi Zingales, 1995, What do we know about capital structure? Some evidence from international data, *Journal of Finance* 50, 1421–1460.

Schaefer, Stephen, and Ilya A. Strebulaev, 2008, Structural models of credit risk are useful: Evidence from hedge ratios on corporate bonds, *Journal of Financial Economics* 90, 1–19.

Smith, Clifford W., and Ross L. Watts, 1992. The investment opportunity set, and corporate financing, dividend, and compensation policies, *Journal of Financial Economics* 32, 262–292.

Strebulaev, Ilya A., 2007. Do tests of capital structure theory mean what they say?, *Journal of Finance 62*, 1747–1789.

Strebulaev, Ilya A., and Baozhong Yang, 2012, The mystery of zero-leverage firms, Working Paper, Stanford University.

Titman, Sheridan, and Roberto Wessels, 1988, The determinants of capital structure, *Journal of Finance* 43, 1–19.

Tserlukevich, Yuri, 2008, Can real options explain financing behavior?, *Journal of Financial Economics* 89, 232–252.

Welch, Ivo, 2004, Capital structure and stock returns, *Journal of Political Economy* 112, 106–131

2.9 Tables and Figures

Table 2.1: Description of Model Parameters and Variables

This table presents a notion index for the single-stage (benchmark) model and the multi-stage model.

Parameter	Definition
r	Risk-free interest rate
μ	Growth rate of cash flows
σ	Volatility of cash flows
τ	Effective corporate tax rate
α	Proportional bankruptcy cost
X_0	Cash flow level (in \$) at time $t = 0$
Π_i, Π	Investment scale of ith stage, $i = 1, 2$
F_i, F	Investment cost (in \$) of ith stage, $i = 1, 2$
C_i, C	Debt coupon (in \$) of ith stage, $i = 1, 2$
X_{Si}, X_S	Investment threshold of ith stage, $i = 1, 2$
X_{Di}, X_D	Default threshold of ith stage, $i = 1, 2$
D_B	Debt value in stage 1 of the benchmark model
D_1	Debt value in stage 1 of the multi-stage model
D_{2i}	Debt value in stage 2 of the multi-stage model, $i = 1, 2$
E_{Bi}	Equity value in stage i of the benchmark model, $i = 0, 1$
E_i	Equity value in stage i of the multi-stage model, $i = 0, 1, 2$
ML_i, ML	Market leverage in ith stage, $i = 1, 2$

Table 2.2: Financing and Investment in Single-Stage and Multi-Stage Models

This table shows the optimal investment and financing decisions of the single-stage benchmark model in Panel A and the multi-stage model in Panel B. The base case parameter values are as follows: risk-free rate $r = 6\%$, growth rate of the cash flow process $\mu = 1\%$, volatility of the cash flow process $\sigma = 25\%$, corporate tax rate $\tau = 10\%$, bankruptcy cost $\alpha = 30\%$, initial value of the cash flow process $X_0 = \$5$, the scales of investment $\Pi_1 = 1$ and $\Pi_2 = 1$, and the investment costs $F_1 = \$100$ and $F_2 = \$200$. The notation index is given in Table 2.1.

Panel A. Single-Stage Model

	$\Pi = 1$	$\Pi = 1$	$\Pi = 1.5$	$\Pi = 0.5$
	$F = \$100$	$F = \$200$	$F = \$100$	$F = \$200$
C	6.522	13.034	6.521	13.034
X_S	12.487	24.973	8.325	49.947
X_D	2.830	5.657	1.887	11.313
ML	0.419	0.419	0.419	0.419

Panel B. Multi-Stage Model

	$\Pi_1 = 1$	$\Pi_1 = 0.75$	$\Pi_1 = 1.25$	$\Pi_1 = 1.5$
	$\Pi_2 = 1$	$\Pi_2 = 1.25$	$\Pi_2 = 0.75$	$\Pi_2 = 0.5$
C_1	5.591	5.447	5.972	6.261
C_2	19.428	14.760	27.205	43.219
X_{S1}	12.364	16.424	9.930	8.300
X_{S2}	23.666	19.359	30.764	44.611
X_{D1}	2.154	2.477	1.959	1.770
X_{D2}	5.429	4.385	7.199	10.737
ML_1	0.283	0.201	0.349	0.390
ML_2	0.378	0.401	0.359	0.343

Table 2.3: Sensitivity of Financing and Investment

This table shows comparative statics of parameters unrelated to the investment process. The base case parameter values are as follows: risk-free rate $r = 6\%$, growth rate of the cash flow process $\mu = 1\%$, volatility of the cash flow process $\sigma = 25\%$, corporate tax rate $\tau = 10\%$, bankruptcy cost $\alpha = 30\%$, initial value of the cash flow process $X_0 = \$5$, the scales of investment $\Pi_1 = 1$ and $\Pi_2 = 1$, and the investment costs $F_1 = \$100$ and $F_2 = \$200$. Relative to the base case, we increase or decrease a parameter by 50% while keeping everything else fixed. The notation index is given in Table 2.1.

	Base	$\tau = 15\%$	$\tau = 5\%$	$\alpha = 45\%$	$\alpha = 15\%$	$\sigma = 37.5\%$	$\sigma = 12.5\%$	$\mu = 1.5\%$	$\mu = 0.5\%$
C_1	5.591	7.427	3.312	4.648	6.766	5.352	5.802	6.084	5.197
C_2	19.428	24.580	12.265	15.603	25.630	25.130	17.364	21.424	17.822
X_{S1}	12.364	12.640	12.004	12.450	12.223	17.509	8.498	11.989	12.745
X_{S2}	23.666	23.677	23.601	24.233	22.521	33.083	16.456	22.908	24.441
X_{D1}	2.154	2.781	1.324	1.819	2.553	1.360	3.447	2.099	2.201
X_{D2}	5.429	6.945	3.380	4.394	7.030	4.711	7.187	5.517	5.348
ML_1	0.283	0.336	0.198	0.250	0.319	0.153	0.533	0.271	0.296
ML_2	0.378	0.445	0.268	0.324	0.451	0.290	0.528	0.380	0.376

Table 2.4: Sensitivity of Model Moments

This table presents the sensitivities of the moments used in the Simulated Method of Moments (SMM) estimation. The structural model parameters that we fit by SMM are the investment scales, Π_1 and Π_2, the bankruptcy cost, α, and the tax rate, τ. The other parameters assume the base case values from Table 2.2. Column 1 presents the model moments for the base case (i.e., $\Pi_1 = 1$, $\Pi_2 = 1$, $\alpha = 0.3$, and $\tau = 0.1$). In columns 2 to 5, each parameter is increased by 50% while keeping the others fixed. The following five moments are used in the SMM. The average quasi-market leverage, QML, is obtained by first calculating cross-sectional averages of quasi-market leverage ratios for every time t and then averaging across time. Quasi-market leverage is defined as the book value of debt divided by the sum of market value of equity and book value of debt. The dispersion of quasi-market leverage ratios, $DispQML$, is the cross-sectional average of the time-series standard deviations of firms' quasi-market leverage ratios. D/K denotes net debt issuance normalized by capital. D/K is calculated only at the investment points. Q is the average market-to-book ratio. Similar to QML, the average is taken first across firms and then across time. Inv/Eq is the average of investment expenditure scaled by the book value of equity at investment points. Panel A displays the sensitivities of the model moments in terms of their absolute changes, while Panel B displays their changes relative to the base case values in the first column of Panel A.

Panel A. Absolute Changes

	Base	$\Pi_1 = 1.5$	$\Pi_2 = 1.5$	$\alpha = 45\%$	$\tau = 15\%$
QML	0.064	0.115	0.064	0.067	0.057
$DispQML$	0.070	0.107	0.070	0.072	0.065
D/K	0.267	0.305	0.231	0.221	0.337
Q	1.184	1.335	1.678	1.190	1.139
Inv/Eq	0.413	0.427	0.394	0.405	0.432

Panel B. Relative Changes

	$\Pi_1 = 1.5$	$\Pi_2 = 1.5$	$\alpha = 45\%$	$\tau = 15\%$
QML	0.805	0.008	0.057	-0.105
$DispQML$	0.527	0.001	0.027	-0.064
D/K	0.140	-0.135	-0.172	0.260
Q	0.128	0.418	0.005	-0.038
Inv/Eq	0.033	-0.048	-0.020	0.046

Table 2.5: Estimation of Model Parameters with Simulated Method of Moments

This table presents the estimation results of the model parameters via Simulated Method of Moments (SMM). The structural model parameters that we fit by SMM are the investment scales, Π_1 and Π_2, the bankruptcy cost, α, and the tax rate, τ. The other parameters assume the base case values from Table 2.2. The following five moments are used in the SMM. The average quasi-market leverage, QML, is obtained by first calculating cross-sectional averages of quasi-market leverage ratios for every time t and then averaging across time. Quasi-market leverage is defined as the book value of debt divided by the sum of market value of equity and book value of debt. The dispersion of quasi-market leverage ratios, $DispQML$, is the cross-sectional average of the time-series standard deviations of firms' quasi-market leverage ratios. D/K denotes net debt issuance normalized by capital. D/K is calculated only at the investment points. Q is the average market-to-book ratio. Similar to QML, the average is taken first across firms and then across time. Inv/Eq is the average of investment expenditure scaled by the book value of equity at investment points. The data moments are calculated using COMPUSTAT's annual tapes for the 1965–2009 period. Four sets of data moments are obtained by using the full sample and three subsamples, which are generated by the tercile cutoffs of Q. Panel A presents the fitted model parameters. The numbers in parentheses are the standard deviation of the fitted parameters b^* across the iterations of SMM with the exception of the χ^2 column in Panel A, in which the numbers in parentheses are the p-values for the overidentification test. Panel B presents the target and fitted moments for each sample.

Panel A. Parameter Estimates

		Π_1	Π_2	α	τ	χ^2
Full Sample	b^*	1.966	1.286	0.324	0.043	0.021
		(0.104)	(0.095)	(0.055)	(0.012)	(0.111)
Low Q	b^*	3.036	0.531	0.440	0.039	0.034
		(0.503)	(0.389)	(0.018)	(0.003)	(0.146)
Medium Q	b^*	2.032	0.848	0.267	0.047	0.036
		(0.134)	(0.144)	(0.091)	(0.015)	(0.133)
High Q	b^*	1.264	2.887	0.284	0.071	0.035
		(0.079)	(0.402)	(0.104)	(0.030)	(0.132)

Panel B. Fitted and Data Moments

		QML	$DispQML$	D/K	Q	Inv/Eq
Full Sample	Data	0.199	0.183	0.166	1.685	0.366
	Fitted	0.198	0.157	0.166	1.761	0.412
Low Q	Data	0.290	0.198	0.126	0.905	0.336
	Fitted	0.284	0.207	0.134	1.054	0.407
Medium Q	Data	0.216	0.169	0.210	1.335	0.373
	Fitted	0.220	0.165	0.224	1.403	0.420
High Q	Data	0.087	0.108	0.176	2.839	0.390
	Fitted	0.086	0.086	0.189	2.731	0.394

Table 2.6: Parameters for Simulation and Descriptive Statistics of Simulated Data

This table presents the parameter values and distributions used for the simulation in Panel A and the descriptive statistics of the simulated leverage ratios in Panel B. To add heterogeneity to the simulated data, several model parameters are randomized at time 0 and kept fixed over time: the investment costs, F_1 and F_2, the bankruptcy cost, α, and the tax rate, τ. In addition, to allow for a correlation structure, the volatility of cash flows is decomposed into a systematic part, σ_S, and an idiosyncratic part, σ_I. β measures a firm's exposure to systematic risk. The investment scales, Π_1 and Π_2, and the means of the bankruptcy cost, α, and the tax rate, τ, are set to the full sample SMM estimates for b^* in Table 2.5. The other parameters assume the base case values from Table 2.2. Panel B reports the distributions of market leverage (ML) and quasi-market leverage (QML). Investment points (Inv. Pts.) refers to the data points where firms are at their investment points. Investment points are further classified as the first and second investment points because there are two stages in the model. All other statistics are given for all data points (i.e. in dynamics). The market leverage ratio, ML, is defined as the market value of debt over the sum of market value of debt and market value of equity, and the quasi-market leverage ratio, QML, is the book value of debt over the sum of market value of equity and book value of debt. For each leverage ratio, the mean, the 1st, 5th, 10th, 50th, 90th, 95th, 99th percentiles, and the standard deviation are reported. For each data set, the statistics are first calculated for each quarter, then averaged across quarters, and then averaged across simulated data sets. Min. and Max. give the minimum and maximum of statistics over the 1,000 data sets. The statistics are based on 1,000 simulated economies, which each contain 400 quarters (after dropping the first 400 quarters) for 3,000 firms.

Panel A. Model Parameters for Simulation

Parameter	Value	Parameter	Distribution
Π_1	2.093	F_1	Uniform[80, 120]
Π_2	1.304	F_2	Uniform[160, 240]
σ_S	0.148	α	Uniform[0.362, 0.542]
N	3,000	τ	Uniform[0.022, 0.032]
Δt	0.25	β	Uniform[0.179, 1.807]
T	100	σ_I	$0.05 + \frac{1}{30}\chi^2(5)$

Panel B. Descriptive Statistics for Leverage

		Percentiles							
	Mean	1%	5%	10%	50%	90%	95%	99%	Std. Dev.
ML									
Inv. Pts.	0.187	0.023	0.046	0.067	0.175	0.323	0.368	0.462	0.099
1st Inv. Pts.	0.162	0.021	0.039	0.057	0.152	0.283	0.318	0.386	0.086
2nd Inv. Pts.	0.237	0.042	0.080	0.109	0.225	0.380	0.426	0.521	0.106
Avg.	0.197	0.000	0.000	0.000	0.089	0.571	0.707	0.899	0.241
Min.	0.113	0.000	0.000	0.000	0.000	0.362	0.503	0.784	0.174
Max.	0.255	0.000	0.000	0.000	0.181	0.698	0.833	0.960	0.285
QML									
Avg.	0.205	0.000	0.000	0.000	0.088	0.606	0.752	0.933	0.254
Min.	0.116	0.000	0.000	0.000	0.000	0.378	0.538	0.838	0.185
Max.	0.266	0.000	0.000	0.000	0.182	0.747	0.876	0.977	0.299

Table 2.7: Capital Structure Regressions on Simulated Data

This table reports average coefficient estimates and average t-statistics (in parentheses) of four cross-sectional regressions over the 1,000 simulated panel data sets from Table 2.6. That is, the regressions are based on 1,000 simulated economies, which each contain 400 quarters (after dropping the first 400 quarters) for 3,000 firms. The regression model is as follows:

$$QML_i = \beta_0 + \beta_1 \, x_i + \beta_2 \, \sigma_i + \beta_3 \, \alpha_i + \beta_4 \, \tau_i + \beta_4 \, \varphi_i + \epsilon_i, \tag{2.26}$$

where x is either profitability, π, in Panel A or market-to-book, Q, in Panel B. In Panel C, we include both profitability, π, and market-to-book, Q, as regressors. We measure profitability, π, as earnings before interest and tax (or cash flows) scaled by total assets, whereas the market-to-book ratio, Q, is the ratio of total market value of asset over book value of asset. The other independent variables are constant firm attributes. They include volatility of cash flows, σ, bankruptcy cost, α, tax rate, τ, and firm size, φ, which equals the sum of book value of debt and book value of equity. The first column (Inv. Pts.) shows OLS regression results using data at investment points only. The regressions in the other columns are for all data points (i.e. in dynamics). The second column (BJK) reports OLS regression results in the fashion of Bradley, Jarrel, and Kim (1984). The dependent variable, QML_i, is calculated as the sum of book values of debt over the 400 quarters divided by the sum of book values of debt and market values of equity over the 400 quarters. The independent variables are calculated similarly (if possible). This definition implies that the dependent variable and independent variables are contemporaneous. The third column (RZ) follows Rajan and Zingales (1995), who define all independent variables as averages over quarters $t-1$ to $t-4$. In this version, the dependent variable QML_i is measured at time t. The last column (FF) adopts the Fama-MacBeth regression approach as in Fama and French (2002). At each time t, QML is regressed on lagged independent variables. Then the time-series of the coefficient estimates are averaged and the standard errors are corrected using the Newey-West method with six lags.

Panel A. Profitability

	Inv. Pts.	BJK	RZ	FF
Constant	0.293	0.371	0.323	0.364
	(18.67)	(12.20)	(6.40)	(47.45)
π	2.208	-0.053	-0.004	-0.012
	(10.25)	(-11.04)	(-5.46)	(-11.83)
σ	-0.791	-0.779	-0.754	-0.843
	(-52.55)	(-30.94)	(-18.03)	(-48.66)
α	-0.328	-0.176	-0.097	-0.107
	(-14.43)	(-2.77)	(-0.94)	(-10.57)
τ	2.785	1.427	0.707	0.775
	(16.40)	(3.00)	(0.92)	(10.06)
φ	0.028	0.067	0.184	0.182
	(7.45)	(8.14)	(18.93)	(27.62)
R^2	0.804	0.313	0.229	0.224
N	2,637	3,000	3,000	1,197,000

Table 2.7 (cont.)

Panel B. Tobin's Q

	Inv. Pts.	BJK	RZ	FF
Constant	0.369	0.403	0.430	0.471
	(38.06)	(14.10)	(9.02)	(59.55)
Q	0.026	-0.049	-0.076	-0.084
	(28.90)	(-20.17)	(-20.23)	(-32.85)
σ	-0.945	-0.937	-1.072	-1.152
	(-112.52)	(-37.49)	(-25.65)	(-66.97)
α	-0.338	-0.208	-0.128	-0.136
	(-16.67)	(-3.49)	(-1.31)	(-14.37)
τ	2.892	1.679	0.916	0.966
	(19.16)	(3.77)	(1.26)	(13.58)
φ	0.033	0.167	0.265	0.259
	(9.93)	(19.96)	(27.95)	(36.20)
R^2	0.844	0.405	0.344	0.233
N	2,637	3,000	3,000	1,197,000

Panel C. Profitability and Tobin's Q

	Inv. Pts.	BJK	RZ	FF
Constant	0.212	0.423	0.444	0.485
	(15.20)	(15.21)	(9.39)	(59.39)
π	3.162	-0.058	-0.005	-0.015
	(16.43)	(-13.08)	(-7.34)	(-11.96)
Q	0.025	-0.052	-0.079	-0.088
	(25.46)	(-21.57)	(-20.95)	(-33.27)
σ	-0.805	-0.958	-1.088	-1.161
	(-58.13)	(-39.36)	(-26.25)	(-67.00)
α	-0.333	-0.208	-0.129	-0.136
	(-17.70)	(-3.59)	(-1.33)	(-14.44)
τ	2.752	1.681	0.924	0.980
	(19.57)	(3.89)	(1.28)	(13.76)
φ	0.065	0.157	0.259	0.252
	(9.28)	(19.11)	(27.48)	(36.01)
R^2	0.866	0.439	0.357	0.240
N	2,637	3,000	3,000	1,197,000

Figure 2.1: Initial Equity Values as a Function of Debt Coupon Choices

This figure charts the intertemporal effect by mapping debt coupon pairs (C_1, C_2) into initial equity value, E_0, on the basis of equation (16). C_1 varies from \$4 to \$7, and C_2 from \$16 to \$22. The base case parameter values are as follows: risk-free rate $r = 6\%$, growth rate of the cash flow process $\mu = 1\%$, volatility of the cash flow process $\sigma = 25\%$, corporate tax rate $\tau = 10\%$, bankruptcy cost $\alpha = 30\%$, initial value of the cash flow process $X_0 = \$5$, the scales of investment $\Pi_1 = 1$ and $\Pi_2 = 1$, and the investment costs $F_1 = \$100$ and $F_2 = \$200$.

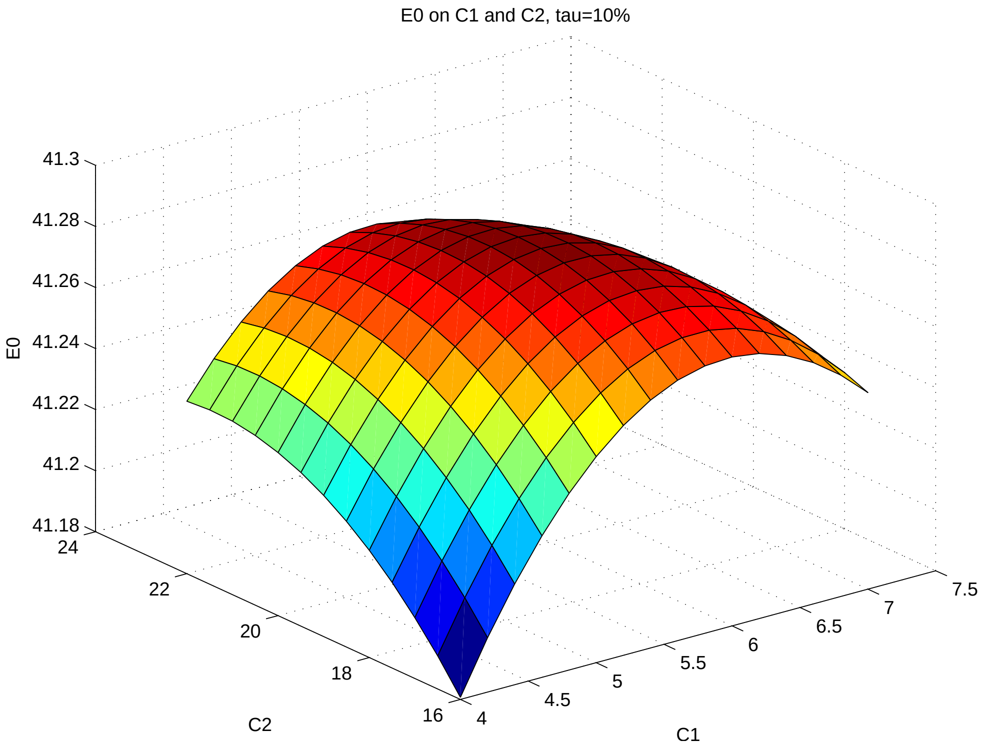


Figure 2.2: Average Leverage of Actual Leverage Portfolios in Event Time

The sample consists of 3,000 firms over 39 years in 1,000 simulated economies based on the extended multi-stage model in Appendix A with three industries, which have 1,000 firms each and are defined based on the subsample (i.e. low, medium, high Q) estimation results for b^* in Table 2.5. The modeling of firm-level heterogeneity follows the procedure in Table 2.6, except that we use here the three subsample estimation results for b^* in Table 2.5. While the initial investment scale is normalized to one, $\Pi_0 = 1$, firms have an exogenously assigned initial coupon, C_0, which is drawn from a lognormal distribution: $C_0 \sim \text{LogNormal}(0.5, 1)$. Each panel presents the average leverage of four portfolios in event time (i.e. quarters), where event time zero is the portfolio formation period. That is, for each quarter in the simulated economies, we form four portfolios by ranking firms based on their actual leverage. Holding the portfolios fixed for the next 20 years, we compute the average leverage for each portfolio. We repeat this process of sorting and averaging for every quarter in our simulated economies. After performing this sorting and averaging for each quarter from quarter 0 to quarter 156, we then average the average leverages across "event time" in each of the simulated economies and then average them across the 1,000 simulated economies to obtain the lines in the figure. The results for book and quasi-market leverage are presented in Panels A and C, where book (quasi-market) leverage is defined as the ratio of book value of debt to book value of assets (sum of book value of debt and market value of equity). Panels B and D present similar results for book and quasi-market leverage, respectively, but for a subsample of firms required to exist for at least 80 quarters (consequently, we can only perform the portfolio formation through quarter 76 for this sample).

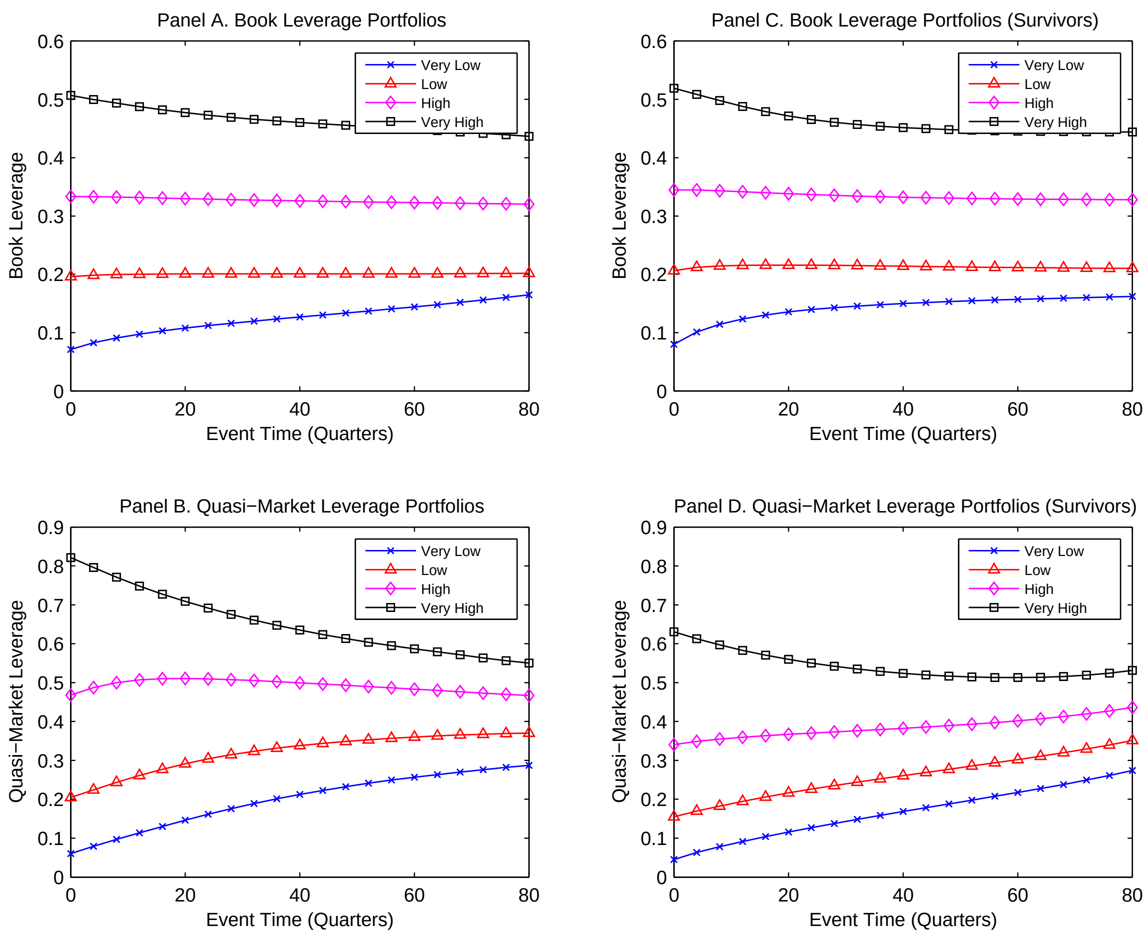
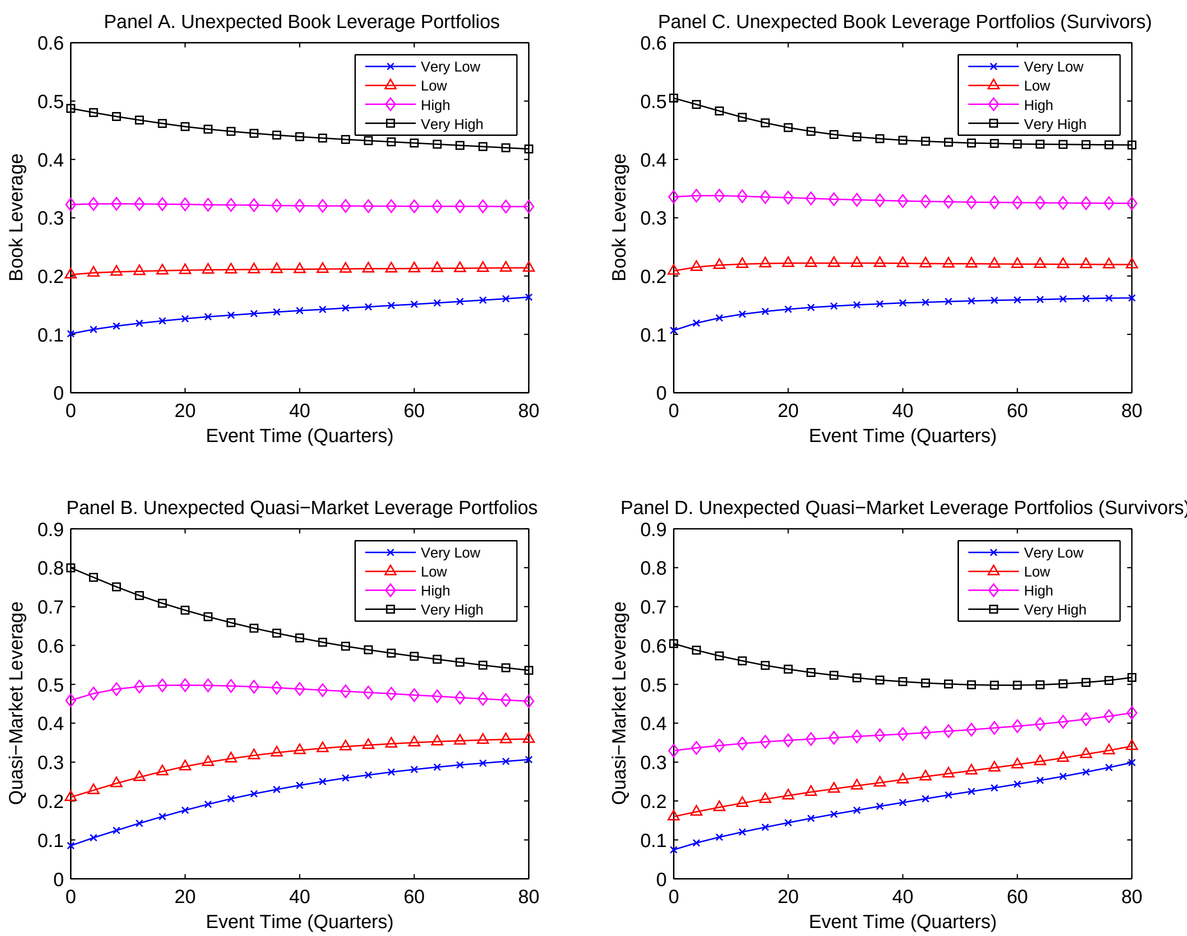


Figure 2.3: Average Leverage of Unexpected Leverage Portfolios in Event Time.

The sample consists of 3,000 firms over 39 years in 1,000 simulated economies based on the extended multi-stage model in Appendix A with three industries, which have 1,000 firms each and are defined based on the subsample (i.e. low, medium, high Q) estimation results for b^* in Table 2.5. The modeling of firm-level heterogeneity follows the procedure in Table 2.6, except that we use here the three subsample estimation results for b^* in Table 2.5. While the initial investment scale is normalized to one, $\Pi_0 = 1$, firms have an exogenously assigned initial coupon, C_0, which is drawn from a lognormal distribution: $C_0 \sim \text{LogNormal}(0.5, 1)$. Each panel presents the average leverage of four portfolios in event time (i.e. quarters), where event time zero is the portfolio formation period. That is, for each quarter in the simulated economies, we form four portfolios by ranking firms based on their unexpected leverage (defined below). Holding the portfolios fixed for the next 20 years, we compute the average leverage for each portfolio. We repeat this process of sorting and averaging for every quarter in our simulated economies. After performing this sorting and averaging for each quarter from quarter 0 to quarter 156, we then average the average leverages across "event time" in each of the simulated economies and then average them across the 1,000 simulated economies to obtain the lines in the figure. The results for book and quasi-market leverage are presented in Panels A and C, where book (quasi-market) leverage is defined as the ratio of book value of debt to book value of assets (sum of book value of debt and market value of equity). Panels B and D present similar results for book and quasi-market leverage, respectively, but for a subsample of firms required to exist for at least 80 quarters (consequently, we can only perform the portfolio formation through quarter 76 for this sample). Unexpected leverage is defined as the residuals from a cross-sectional regression of leverage on market-to-book, Q, profitability, π, volatility of cash flows, σ, bankruptcy cost, α, tax rate, τ, firm size, φ, and industry indicator variables, where all independent variables are lagged one year.



Chapter 3

Does Idiosyncratic Risk Matter for the Cross-section of Stock Returns? Evidence from the Option Markets

3.1 Introduction

This paper re-examines whether idiosyncratic risk explains the cross-section of stock returns using implied idiosyncratic variance extracted from option prices. Starting from the mean-variance analysis by Markowitz (1959), asset pricing models such as the Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965) and Black (1972), the Ross (1976) arbitrage pricing theorem (APT), and the Fama-French (1993) three-factor model establish a framework in which idiosyncratic risk does not explain stock returns because it can be eliminated through diversification. However, if investors can not hold fully diversified portfolios, they will care about idiosyncratic risks and request compensation for bearing it. Based on this argument, Merton (1987) provides a theoretical model implying that the conditional expected stock return is positively related to its own idiosyncratic variance, i.e. firm-specific component of the firm's return variance.

Enlightened by Merton (1987)'s model, this paper re-examines the linear relationship between expected stock returns and idiosyncratic risk by studying idiosyncratic variance instead of volatility. This has never been done in the literature. Moreover, the majority of the studies in this literature use past realized idiosyncratic volatility as a proxy for expected idiosyncratic volatility, and find conflicting relationships. This paper adopts the implied idiosyncratic variances extracted from option prices as the measure of expected idiosyncratic risk. This measure is also used in Diavatopoulos, Doran and Peterson (2008) (DDP hereafter). According to DDP, implied idiosyncratic volatility is a better measure for expected idiosyncratic volatility than past realized idiosyncratic volatility, but their tests fail to provide support to this claim. They also document a positive relationship between implied volatility and future stock returns, which is consistent with Merton (1987)'s prediction and poses a challenge to the standard asset pricing models. This paper differs from DDP in many aspects. First, it focuses on reconciling the conflicting empirical findings in the literature by decomposing realized idiosyncratic variance into expected (implied) and unexpected components. Second, the sample in this paper covers stocks with tradable options during an extended period of January 1996 to December 2006, compared to January 1996 to June 2005 in DDP. Third, it provides a formal test of the predictive power of implied idiosyncratic variance on future realized idiosyncratic variance and shows further evidence that implied idiosyncratic variance is a better predictor than historical idiosyncratic variance. Last but not least, DDP uses the implied volatilities of standardized options in OptionMetrics directly, whereas in this paper, the implied idiosyncratic variances are extracted from the implied volatilities on one-month-to-expiration at-the-money options. The stock holding period returns are calculated corresponding to the terms of these options, i.e. from immediately after the option expiration Saturday in calendar month t to option expiration date in calendar month $t+1$. There are two advantages for this approach. First, the underlying stocks of these options are the worst candidates for short-sale constraints, which set a proper setting to test Merton (1987) model (frictionless market). Second, implied volatilities for these options tend to be a better measure of future idiosyncratic volatilities than other measures.

To study the effect of implied, unexpected and realized idiosyncratic variances on stock returns, Fama-MacBeth (1973) regressions are carried out for the full sample, DDP's sample period as well as NYSE and NYSE/AMEX sub-samples. The tests imply a significant positive effect of implied idiosyncratic variance for low idiosyncratic variance sample, a significant negative effect of past realized idiosyncratic variance for high idiosyncratic variance sample and no significant effect for a sample with medium level of idiosyncratic variance. Furthermore, the positive relationship in DDP does not hold in the extended sample period. More specifically, no significant relationship exists between implied idiosyncratic variance and future stock returns for the period of January 1996 through December 2006. Instead, past realized idiosyncratic variance shows a significant negative effect on future stock returns for both the extended sample period and the sample period in DDP. This second finding is consistent with Ang, Hodrick, Xing and Zhang (2006a) (AHXZ hereafter). What is more interesting and confusing lies in the results of sub-sample analyses. The positive relationship found in DDP resurfaces for NYSE stocks sub-sample. Neither implied nor past realized idiosyncratic variances is significant for NYSE/AMEX stocks sample. Therefore there is no single robust relationship between implied (realized) idiosyncratic variance and future stock returns. In an attempt to resolve the puzzling findings, I show that the negative effect of past realized idiosyncratic variance is driven by the persistent idiosyncratic variance shock and it is the most significant for high idiosyncratic variance stocks. The positive effect of implied idiosyncratic variance is the most significant for low idiosyncratic variance stocks. The tradeoff between these two contradictive effects can yield negative, positive and even no relationships between idiosyncratic risk and cross-sectional stock returns. Therefore the mixed findings in the literature can be contributed to this tradeoff.

The rest of the paper is organized as follows: Section 3.2 provides a review of the related literature on the idiosyncratic risk-stock returns relationship and the information content in implied volatility. Section 3.3 defines idiosyncratic risk measures, presents descriptive statistics of the samples, and tests the predictive power of implied idiosyncratic variance for future realized idiosyncratic variance. Section 3.4 re-examines the relationship between implied (past) idiosyncratic variance and future stock returns in DDP's sample periods and another two sub-samples. Section 3.5 briefly discusses the cause of the mixed results and section 3.6 concludes.

3.2 Related Literature

In addition to Diavatopoulos, Doran and Peterson (2008) with pieces of evidence from the options markets, there is a strand of research debating on whether idiosyncratic risk matters for expected stock returns using stock market information.

Theoretically Merton (1987) indicates that if investors are not able to hold the market portfolio, firms with larger idiosyncratic variance have higher returns to compensate the investors for bearing the idiosyncratic risk. Barberis and Huang (2001) also develop an intertemporal behavior model that implies a positive relation between idiosyncratic risk and stock returns because of investors' loss aversion and mental accounting on stock level.

Supporting the theoretical models, Malkiel and Xu (2004) find a significantly positive explanatory power of idiosyncratic volatility for cross-sectional expected stock returns at the portfolio level. Goyal and Santa-Clara (2003) use time-series analysis to identify a significant positive relationship between average total (85% idiosyncratic) volatility and the return of the S&P 500 index market. Chua et al (2006) decompose idiosyncratic volatility into expected and unexpected components using AR (2) model and find that expected (unexpected) stock returns is positively associated with expected (unexpected) idiosyncratic volatilities. Fu (2009) and Spiegel and Wang (2005) consider the time-varying property of the conditional idiosyncratic volatilities. Using monthly data, they estimate the conditional idiosyncratic volatilities with more advanced models such as GARCH and EGARCH and find significantly positive relation.

On the contrary, Bali and Cakici (2006) argue that the empirical evidence is not robust to different choices of data frequency, weighting scheme, conditioning variables, and sample periods. Wei and Zhang (2005) revisit the study by Goyal and Santa-Clara (2003) with longer sample periods but find no sustained results. Ang, Hodrick, Xing and Zhang (2006a)(AHXZ hereafter) measure idiosyncratic volatility of individual stocks relative to Fama-French (1993) three-factor model and then build portfolios by sorting on stock's idiosyncratic volatility. Contrary to all of the above findings, AHXZ show a strong negative relationship and impose a puzzle to this literature.

The main contribution of DDP is the use of implied idiosyncratic volatility extracted from option market information as the measure for expected idiosyncratic volatility. Canina and Figlewski (1993) examine S&P 100 index options (OEX) from 1983 to 1987 and conclude that implied volatility has no correlation with future volatility. Christensen and Prabhala (1998) use a longer data period, correct for biases related to measurement error and find that implied volatility of S&P 100 index options outperforms past volatility in forecasting future index volatility. Granger and Poon (2003) survey the literature on forecasting the volatility and conclude that the market forecast embedded in implied volatility is the best forecast of future realized volatility. Goyal and Saretto (2006) investigate the predictability of the cross-section of individual equity option implied volatilities and conclude that the cross-section of stock implied volatilities lead to better predictions of future volatility than those provided by the market's implied volatility. Giot (2005) finds that VIX is useful for predicting returns on various indices formed on the basis of size and growth versus value characteristics. Banerjee, Doran, and Peterson (2007) find that VIX levels and innovations predict the returns of characteristic based portfolios.

3.3 The Idiosyncratic Volatility Measures and the Data

Instead of using calendar months as holding periods of stocks, I use the one month period from the first trading day (usually Monday) immediately after the option expiration date (usually Saturday) in month t to the option expiration date in month $t+1$. Throughout this paper I will refer to months as the definition here and calendar months as regular months. Defining holding period in this way has a few advantages. First, one-month-to-expiration at-the-money options are the most liquid option contracts. Hence the underlying stocks of these options are least likely subject to short-sale constraints. Second, implied volatility of at-the-money options is directly related to the underlying volatility. It carries similar statistical properties as it is very persistent. Therefore it is a good measure of market's expectation on the underlying volatility and is superior to past realized volatility.

3.3.1 The Measures of Idiosyncratic Variances

Idiosyncratic Variance

Fama and French (1993) three-factor model as shown below has been one of the most successful factor-based asset pricing models.

$$r_t^i = \alpha^i + \beta^i_{MKT} MKT_t + \beta^i_{SMB} SMB_t + \beta^i_{HML} HML_t + \epsilon^i_t, \tag{3.1}$$

where MKT is the market excess return measured by subtracting the one-month T-bill rate from the market return, SMB is the return difference between small size stock portfolio and big size stock portfolio. HML is the return on high book-to-market stock portfolio minus the return on low book-to-market stock portfolio. Daily MKT, SMB and HML are provided by Kenneth French on his website[1].

By using Fama-French three-factor model (3.1) to describe the cross-section of stock returns, I am able to extract the idiosyncratic variance from a stock's total variance by following the approach in Ang et al (2006a). At the end of each month, for every stock i with options trading on it, daily stock returns over this month are regressed on MKT, SMB and HML. The idiosyncratic variance for stock i at t (IV_t^i) is defined as the variance of the residuals in model (3.1), i.e.

$$IV_t^i = Var(\epsilon_t^i). \tag{3.2}$$

IV_t^i is then annualized by multiplying $\frac{252}{N}$, where N is the number of trading days in that month. This measure is based on historical realized returns, hence is the realized idiosyncratic variance.

Idiosyncratic variance can also be defined as the annualized variance of residuals from the CAPM. I prefer to using Fama-French three-factor model because it is more powerful in capturing the cross-sectional variations in stock returns (Fama and French (1993)). Given that CAPM fails to explain the cross-section of stock returns for the recent period (Fama and French (1992a)), idiosyncratic variance measured relative to Fama-French three-factor model is more appropriate.

Implied Idiosyncratic Variance

Implied idiosyncratic variance (IIV_t^i) is the idiosyncratic component of implied variance. As is the argument for implied volatility, if option market is efficient, implied idiosyncratic variance should reflect market investors' expectation on future idiosyncratic variance of the underlying stock returns. Without the availability of implied volatilities on SMB and HML portfolios, I use the market model (3.3) to express the cross sectional stock return processes for decomposing implied variance.

$$r_t^i = \alpha^i + \beta^i MKT_t + \epsilon_t^i, \tag{3.3}$$

where MKT_t is the excess return on the market and β^i is defined as stock i's Beta. Taking variances on both sides, we can see that the implied idiosyncratic variance of stock i is the difference between its implied variance and the implied variance of the market scaled by Beta-squared.

$$IIV_t^i = ImpVar_t^i - (\beta^i)^2 ImpVar_t^{MKT}, \tag{3.4}$$

[1] $http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\_library.html$

where $ImpVar_t^i$ and $ImpVar_{MKT}^i$ are the implied variances for stock i and the market at the end of month t.

The Chicago Board Options Exchange Volatility Index (VIX) is a popular measure of the 30-day implied volatility of S&P 500 index options. It is annualized and in percentage term. In this paper I use VIX as a proxy for the implied volatility of the market. Hence $ImpVar_t^{MKT} = (\frac{VIX_t}{100})^2$. $ImpVar_t^i$ is the square of the average of implied volatilities for at-the-money put and call options with one month to expiration. β^i is estimated with the market model 3.3 using previous 90 daily returns. Therefore stocks in my sample are required to have at least 91 consecutive trading days. Theoretically implied idiosyncratic variance is non-negative. But it could be negative if we calculate it with equation (3.4). So I set IIV_t^i to zero if this happens. Given that the proportion of such data is really low in the sample (less than 5%) this adjustment should not have a significant effect empirically.

Unexpected Idiosyncratic Variance

The difference between the actual realized idiosyncratic variance for month $t+1$ and the implied idiosyncratic variance estimated at the beginning of month t is called unexpected idiosyncratic variance at $t+1$ (UIV_{t+1}). This is a measure of idiosyncratic variance shock and is known to the investors only after idiosyncratic variance is realized.

$$UIV_{t+1} = IV_{t+1} - IIV_t. \tag{3.5}$$

Although the definition is similar, UIV_{t+1} is not the idiosyncratic volatility risk premium. The implied idiosyncratic variance is biased and is not the idiosyncratic component of the risk-neutral implied volatility. It can be thought of as the idiosyncratic variance realized out of investors' expectations.

3.3.2 The Data and The Descriptive Statistics

The Data

The daily implied volatility data for this paper is obtained from the OptionMetrics Ivy database for the period of January 1996 to December 2006. Daily stock prices, returns, number of shares outstanding are provided by CRSP and book value of equity can be obtained from COMPUSTAT. I download daily VIX levels calculated with the new methodology from Chicago Board Options Exchange (CBOE) website[2] and daily returns for Fama-French factors (MKT, SMB and HML) from Kenneth French's website.

The sample includes stocks with tradable options during the period of January 1996 to December 2006. I restrict the sample to stocks traded on New York Stock Exchange (NYSE), American Stock Exchange (AMEX) and NASDAQ. I also require that the firms in the sample to have at least 90 previous consecutive daily stock returns and implied volatilities are available on the first trading day after option expiration date in a month. This leaves me with an overall sample of 3185 firms.

The Descriptive Statistics

For my analysis in Section 3.4, I use four samples: (1) the complete sample covering the period of January 1996 to December 2006; (2) DDP sample: for the sub-period of January 1996 to June

[2]http://www.cboe.com/micro/vix/historical.aspx

2005; (3) NYSE and AMEX stocks sub-sample; (4) NYSE stocks sub-sample. All the ETFs, foreign, financial and utility stocks are excluded from sample (2) to (4). Panel A to E in Table 3.1 shows the descriptive statistics for these samples. There are 3185 firms in sample (1), 2136 in sample (2), 1241 in sample (3), and 1192 in sample (4). For all of the variables in the table except #firms/month, time-series averages for each firm are taken first, and then descriptive statistics are calculated for the cross-section of these time-series means. For #firms/month, number of firms in every month is calculated first and then the average is taken for these monthly numbers.

$SIZE$ is the market value of equity in thousand of million dollars and is computed by multiplying stock's close price and number of shares outstanding at each option expiration date. Book-to-market equity ratio (B/M) at the end of each month t is the book value of equity at the end of the previous year divided by the market value of equity at the end of month t. Beta is the slope in the market model estimated with previous 90 daily stock returns. RV is the realized variance measured by the variance of daily stock returns in a month. $ImpVar$ is the squared implied volatility for one-month-to-expiration at-the-money options. IV represents the one-month realized idiosyncratic variance. IIV refers to the one-month implied idiosyncratic variance and UIV the unexpected idiosyncratic variance. All of the variance measures are annualized.

As shown in Panel A, the average (median) size for the complete sample is 4.08 (0.92) thousand of million dollars. This indicates that the sample firms are on average large firms and are less likely to make short sale constraints binding. About 50% of the sample firms are low book-to-market firms. The median firm has a B/M equal to 0.4. The sample provides a good spread for Betas. Half of the firms are less risky than the market and the other half riskier. As to the variances, the average firm has an implied variance (volatility) of 31.80% (56.39%), higher than the mean realized variance. Realized variance is on average 27.84%, 89% of which is the idiosyncratic component. This is similar to Goyal and Santa-Clara (2003)'s finding. The average realized idiosyncratic variance (24.95%) is only 0.43% higher than the average implied idiosyncratic variance (24.51%), implying that implied idiosyncratic variance is a good proxy for realized idiosyncratic variance. We can also observe that on average investors receive a positive shock of 0.43% to their expected idiosyncratic variance. So the investors tend to underestimate idiosyncratic variance on average. Note that the shocks can be both positive and negative.

Comparing Panel A through E, we can see that all four samples are similar in B/M and Beta. The NYSE sample contains the largest stocks and the complete sample the smallest. DDP sample is the most volatile sample. Between the NYSE/AMEX sample and the NYSE sample, the firm size for NYSE sample is 260 millions larger on average and the variances are about 1.5% lower except for UIV. The characteristics of my DDP sample do not match DDP's. Their sample has larger size, lower volatilities. This sample difference is caused by our different filters for stocks. They keep stocks other than those traded on NYSE/AMEX/NASDAQ and require at least five years of trading history. So their sample filters are in favor of large established firms. Giving a thorough consideration of all the characteristics of the samples, the NYSE sample is the closest to DDP's sample. Therefore, one can expect to see a positive relationship between implied idiosyncratic variance and future stock returns with this sample.

3.3.3 The Predictive Power of Implied Idiosyncratic Variance

Before jumping into the tests of the relationship of interest, a natural question to ask is: why bother using implied idiosyncratic variance instead of historical idiosyncratic variance in the tests? Aside from the impact of mitigating short sale constraints that options have on their underlying stocks, implied idiosyncratic variances derived from option prices are also a better predictor of future idiosyncratic variances than historical idiosyncratic variance. In this section, I study the predictive power of implied and past realized idiosyncratic volatility to show the second point. The typical regressions for showing this in the literature are given as follows:

$$IV_{t+1} = a_{iiv} + b_{iiv} IIV_t + e_{t+1}, \tag{3.6}$$

$$IV_{t+1} = \alpha + \eta IIV_t + \gamma IV_t + e_{t+1}, \tag{3.7}$$

$$IV_{t+1} = a_{iv} + b_{iv} IV_t + e_{t+1}. \tag{3.8}$$

If implied idiosyncratic variance contains information about future realized idiosyncratic variance, we expect to see b_{iiv} significantly positive. Moreover, if the joint test of $a = 0$ and $b_t = 1$ is rejected, IIV_t is biased and if $\gamma > 0$ is supported, it is informationally inefficient.

The tests take two steps. As the first step, I run regressions 3.6, 3.7 and 3.8 for each firm in the sample. Then the mean of the coefficient estimates across firms are calculated and inferences are drawn based on these coefficient means.

Table 3.2 presents the coefficient estimates and t-statistics for the above regressions. Both implied and past realized idiosyncratic variances have significantly positive predictive power, but when used together both of them are still significant. With coefficient estimates of 0.5804 and 0.3361 which are far from 1, they are biased estimators. Although the results are not reported in the table, the augmented Dickey-Fuller tests support this claim. Therefore, implied and past realized idiosyncratic variances are both biased and inefficient estimator. However given that it explains 17.64% of the variations in future realized idiosyncratic variance, i.e. 3.62% higher than past realized idiosyncratic variance does, implied idiosyncratic variance is a better estimator for future idiosyncratic variance than past realized idiosyncratic variance. Tests using implied idiosyncratic variances are likely to draw a better picture of relationship between idiosyncratic risk and expected stock returns.

My findings here confirm DDP's claim about the predictive power of implied idiosyncratic variance. Instead of using realized idiosyncratic variance directly as the dependent variable, they run the regression of total realized variance on implied idiosyncratic variance with control of market implied variance. Yet they do not report the adjusted R-square.

3.4 Fama-MacBeth (1973) Regressions of Stock Returns on Implied and Past Realized Idiosyncratic Variances

3.4.1 The Methodology

I examine the relationship between cross-sectional stock returns and implied and past realized idiosyncratic variances using Fama-MacBeth (1973) regressions. This is a two-stage procedure. In the first stage, for every month $t + 1$, monthly holding period returns for all stocks are regressed on contemporaneous implied idiosyncratic variances, past realized idiosyncratic variances, firm characteristics and other controls. In the second stage, I calculate the mean of the time-series of the coefficient estimates from the first stage regression for each independent variable, and test if they are significantly different from zero. The regressions in the first stage take the form:

$$\begin{aligned} r^i_{t+1} = {} & \alpha_{t+1} + \beta_{IIV,t+1} IIV^i_t + \beta_{IV,t+1} IV^i_t + \beta_{size,t+1} LnSIZE^i_t \\ & + \beta_{B/M,t+1} LnB/M^i_t + \beta_{r,t+1} r^i_t + \eta_{Beta,t+1} Beta^i + \xi_{t+1} z^i + \epsilon^i_{t+1}, \end{aligned} \tag{3.9}$$

where r^i_{t+1} is stock i's holding period return from t to $t+1$; IIV^i_t is stock i's implied idiosyncratic variance estimated at the end of t for the term from t to $t+1$; IV^i_t is the realized idiosyncratic variance for stock i over the previous month from $t-1$ to t; $LnSIZE^i_t$ is the natural log of firm i's market value of equity at t; LnB/M^i_t is the natural log of firm i's book-to-market equity ratio; r^i_t is the realized holding period return for the previous month from $t-1$ to t, and $Beta^i$ is stock i's Beta in the market model; z^i represents other possible controls. I choose past unexpected idiosyncratic variances UIV^i_t and past expected idiosyncratic variance IIV^i_{t-1} as the extra controls in some of the regressions. This will help us find out the potential driver of the results. Since the stocks are held for one month only, the holding periods are non-overlapping.

In these regressions the coefficients $\beta_{IIV,t+1}$ and $\beta_{IV,t+1}$ should be zero under the null hypothesis that standard asset pricing models hold. I run regressions on a monthly basis with all the returns and variances being annualized. The sign and significance of $\beta_{IIV,t+1}$ and $\beta_{IV,t+1}$ show the relation between implied and past realized idiosyncratic variances and expected stock returns. Merton (1987) model predicts that $\beta_{IIV,t+1} > 0$ and significant.

3.4.2 Regressions in the Complete Sample

Table 3.3 reports the coefficient estimates of the Fama-MacBeth (1973) cross-sectional regressions with t-statistics reported in parentheses. All regressions have common controls for $LnSIZE^i_t$, LnB/M^i_t, past return r^i_t, and factor loading $Beta^i$, but have variations in adding current implied idiosyncratic variance IIV^i_t, implied idiosyncratic variance for the previous month IIV^i_{t-1}, past realized idiosyncratic variance IV^i_t, and past unexpected idiosyncratic variance UIV^i_t.

As in Fama and French (1993), $SIZE$ and B/M have significant explanatory power to the cross-section of stock returns. Past return has significant negative effect showing strong mean reversion and the factor loading on market is not significant here.

Column (1) to column (3) indicates that $\beta_{IIV,t+1}$ is not significantly different from zero, i.e. IIV^i_t does not explain the cross-section of stock returns for the complete sample period. In the first two regressions, $\beta_{IIV,t+1}$ are slightly negative, contrary to Merton (1987)'s prediction.

On the other hand, $\beta_{IV,t+1}$ is significant at 5% level as shown in column (3) to (5). Conditional on firm characteristics, past return and factor loading, 1% past realized idiosyncratic variance translates to 0.0733% decrease in expected stock return. The findings here are contrary to DDP's main findings, but consistent with AHXZ. Another interesting finding here is that past unexpected idiosyncratic variance, UIV^i_t has a significantly negative effect on future stock returns when past expected idiosyncratic variance is controlled. When UIV^i_t is used together with IV^i_t, it takes away IV^i_t's explanatory power. This suggests that the UIV^i_t might be the driver of the significant negative explanatory power of past realized idiosyncratic variance.

To see if difference in sample period contributes to the contradiction to DDP, I run the regressions using their sample period in the next section.

3.4.3 Regressions for my DDP Sample Period

Table 3.4 shows the coefficient estimates for the same set of regressions as in Table 3.3, but for a shortened sample period of January 1996 to June 2005. This is the same period used in DDP. All the ETFs, foreign stocks, financial and utilities stocks are excluded, following DDP. This leaves me with 2136 firms in total, 117 less than their sample. I restrict my sample to optionable stocks traded on NYSE/AMEX/NASDAQ only while they keep all optionable stocks. This difference in filters causes the difference in our sample sizes and sample characteristics.

From the table you can see that current implied idiosyncratic variance positively explains contemporaneous stock returns. In column (3), 0.0774% stock returns can be explained by 1% contemporaneous implied idiosyncratic variance. The coefficients in this table can not be directly compared to those in DDP because they use volatility measures as opposed to variance measures in this paper. Although the sign of $\beta_{IIV,t+1}$ is consistent with DDP, it is not significant here.

The significant negative effect of past realized idiosyncratic variance found in the previous section is sustained and strengthened for this sample period. Moreover past unexpected idiosyncratic variance is still significantly negatively correlated with future stock returns conditional on past expected implied idiosyncratic variances. 1% shock in past idiosyncratic variance leads to 0.0921% decrease in future stock return.

To this point, it appears to us that implied idiosyncratic variance has no relationship with future stock returns. But historical idiosyncratic variance is negatively correlated with future stock returns, consistent with Ang et al. (2006a). There is no evidence supporting Merton (1987) up to now, but it is not clear whether the result here is robust. In the next two sections, I check the robustness of the findings in NYSE/AMEX sub-sample and NYSE sub-sample.

3.4.4 Regressions for NYSE/AMEX Stocks

To address the concern that the results in Section 3.4.2 and 3.4.3 are driven by small stocks which are more likely to be subject to short sale constraints, I exclude NASDAQ together with all the ETFs, foreign, financial and utility stocks from the complete sample. As shown in Table 3.5, now $\beta_{IIV,t+1}$ is larger, positive and more powerful, but still not significant. The significance of the negative correlation between r_{t+1} and IV_t disappear and the sign of the correlation switches to positive. It seems that neither implied idiosyncratic variance nor past realized idiosyncratic variance matters for the cross-section of stock returns for normal to large size stocks.

3.4.5 Regressions for NYSE Stocks Only

Altogether 1192 firms are used in Table 3.6, about 1/3 of the complete sample. Among all the samples I examine in section 3.4, this sub-sample has the largest stock market values and the lowest idiosyncratic variances on average. When NYSE stocks are considered alone, DDP's finding resurfaces. Implied idiosyncratic variance is significantly positively associated with stock returns at 5% level. 1% implied idiosyncratic variance explains 0.2410% future stock returns. The effect of past realized idiosyncratic variance becomes positive but insignificant.

As discussed in Section 3.3, this sample is the one most similar to the sample used by DDP. So it is not surprising that we find significant positive relationship for this sample. Given the characteristics of this sub-sample, we can conclude that the significant positive relationship in DDP is mainly driven by large size and low volatility stocks.

In summary, in the process of moving from small, high idiosyncratic variance stocks sample to large, low idiosyncratic variances stocks sample, the positive correlation between implied idiosyncratic variance and future stock return strengthens and the negative correlation between past realized idiosyncratic variance and future stock return weakens. So there is no single robust relationship between idiosyncratic variance and stock returns based on both measures. It seems that two conflicting effects related to size and level of idiosyncratic variance co-exist. Therefore what we observe relies on which factor dominates.

3.5 A Possible Explanation

This section serves the purpose of explaining the puzzling finding in Section 3.4. The analysis in the last section show that: (1) small size and high idiosyncratic variance stocks are sensitive to idiosyncratic variance shocks; (2) the idiosyncratic variance shock UIV_t is the driving factor for the significant negative effect that past realized idiosyncratic variance has on future stock returns, However it is not very efficient in that its own power is gone when IV_t is controlled. So there must be a better driving factor; (3) large size and low idiosyncratic variance stocks are more sensitive to the positive effect of implied idiosyncratic variance.

These findings could be explained by a possible behavioral story. It is possible that investors view idiosyncratic variance shocks for high variance stocks and low variance stocks differently. For low idiosyncratic variance stocks, a small shock on average to the idiosyncratic variance would not surprise investors much because even after the shock the idiosyncratic variances are still low and the shock tends to be eliminated through mean-reversion process soon. Thus their expectations on idiosyncratic variance do not change much. So the effect of the expectation dominates past shock. It is a different story for high idiosyncratic variance stocks. They are more volatile and riskier to investors, and usually experience large idiosyncratic shocks. For this type of stocks investors pay close attention to the idiosyncratic variances and shocks especially the type of shocks that persists. Also they require a volatility risk premium. Therefore the effect of past shock dominates expectations.

Table 3.7 reports the coefficient estimates for the following regression with the samples used in Section 3.4.

$$\begin{aligned} r^i_{t+1} \quad = \quad & \alpha_{t+1} + \beta_{IIV,t+1} IIV^i_t + \beta_{IV,t+1} IV^i_t + \beta_{UIV} UIV^i_{t-1} + \beta_{size,t+1} LnSIZE^i_t \\ & + \beta_{B/M,t+1} LnB/M^i_t + \beta_{r,t+1} r^i_t + \eta_{Beta,t+1} Beta^i + \epsilon^i_{t+1}, \end{aligned} \tag{3.10}$$

where UIV^i_{t-1} is the unexpected idiosyncratic variance two months before stock return is realized. It is a proxy for persistent idiosyncratic variance shock. The testable hypotheses for the discussion above are: (1) if β_{IIV} is significantly positive, Merton (1987) is supported; (2) if β_{IV} is not significant while β_{UIV} is significantly negative, the persistent shock UIV_{t-1} is the driving factor of the negative relation between stock returns and past realized idiosyncratic variance.

Looking horizontally at Table 3.7, we can see that (1) β_{IIV} has correct sign according to Merton (1987) in three of the samples and is significant at 5% level in the NYSE sample which contains large size and low volatility stocks. For the complete sample, β_{IIV} is negative but insignificant. This is a distortion caused by the large negative influence of the persistent shock UIV_{t-1}; (2) IV_t has no significant predictive power for stock returns once UIV_{t-1} is controlled. UIV_{t-1} shows strong negative correlation with stock returns at 10% level for the complete model and 5% level for the DDP sub-sample. It is negative but not significant for the other two samples having larger and less volatile stocks. This means that the negative relationship between stock returns and IV_t comes from the effect of unexpected idiosyncratic variance one month before IV_t is realized. This effect dominates when sample stocks are more volatile and is dominated by implied idiosyncratic variance when sample stocks are less volatile. This illustrative analysis is consistent with the story above.

The explanations in this section shed light on the mixed results found in the literature. It is possible that previous studies all find a specific situation caused by the interaction of two factors with opposite effect on future stock returns. This paper tells us what is going on behind the scene.

3.6 Conclusion

I revisit the relationship between idiosyncratic risk and the cross-section of stock returns using information from the options market. By decomposing realized idiosyncratic variance into expected (implied) and unexpected (shock) components, I find that there is a significant positive correlation driven by implied idiosyncratic variance for large size and low volatility stocks, and a significant negative correlation of past realized idiosyncratic variance for future stock returns for small size and high volatility stocks. Moreover, no relationship is found for NYSE/AMEX stocks that have moderate size and volatilities. These results provide an explanation to the mixed findings in the literature. It is likely that investors have different perceptions towards idiosyncratic variance shocks. Persistent idiosyncratic variance shocks affect future stock returns negatively and investors of small size and high idiosyncratic volatility stocks pay more attention to this shock. Therefore, the negative effect dominates. On the other hand, implied idiosyncratic variances has a positive effect and investors of stocks with large size and low idiosyncratic risk pay more attention to this effect. The combined influence of these two conflicting factors could result in positive relationship, negative relationship or none depending on the strength of the two effects. The tests can be extended by controlling for momentum, which is left as a future research.

3.7 References

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006a, The Cross-Section of Volatility and Expected Returns, *Journal of Finance 61*, 259-299.

Bakshi, Gurdip, and Nikunj Kapadia, 2003, Delta-Hedged Gains and the Negative Market Volatility Risk Premium, *Review of Financial Studies 16*, 527-566.

Bali, Turan G., Nusret Cakici, Xuemin Yan, and Zhe Zhang, 2005, Does Idiosyncratic Risk Really Matter? *Journal of Finance 60*, 905-929.

Banerjee, Prithviraj S., James S. Doran, and David R. Peterson, 2007, Implied Volatility and Future Portfolio Returns, *Journal of Banking and Finance 31*, 3183-3199.

Barberis, Nicholas, and Ming Huang, 2001, Mental Accounting, Loss Aversion, and Individual Stock Returns, *Journal of Finance 56*, 1247-1292.

Battalio, Robert, and Paul Schultz, 2006, Options and the Bubble, *Journal of Finance 61*, 2071-2102.

Christensen, B.J., and N.R. Prabhala, 1998, The Relation between Implied and Realized Volatility, *Journal of Financial Economics 50*, 125-150.

Chua, Choong Tze, Jeremy Goh, and Zhe Zhang, 2005, Idiosyncratic Volatility Matters for the Cross-Section of Returns P In More Ways than One! *Working paper, Singapore Management University.*

Copeland, Maggie M., and Thomas E. Copeland, 1999, Market Timing: Style and Size Rotation using the VIX, *Financial Analysts Journal 55*, 73-81.

Danielsen, Bartley R., and Sorin M. Sorescu, 2001, Why Do Option Introductions Depress Stock Prices? An Empirical Study of Diminishing Short-Sale Constraints, *Journal of Financial and Quantitative Analysis 36*, 451-484.

Diavatopoulos, Dean, James S. Doran, and David R. Peterson, 2008, The Information Content in Implied Idiosyncratic Volatility and the Cross-Section of Stock Returns: Evidence from the Option Markets, *The Journal of Futures Markets 28*, 1013-1039.

Dennis, Patrick J., Stewart Mayhew, and Christopher T. Stivers, 2006, Stock Returns, Implied Volatility Innovations, and the Asymmetric Volatility Phenomena, *Journal of Financial and Quantitative Analysis 41*, 381-406.

Doran, James S., and Ehud I. Ronn, 2006, The Bias in Black-Scholes/Black Implied Volatility: An Analysis of Equity and Energy Markets, *Review of Derivatives Research 8*, 177-198.

Douglas, George Warren, 1969, Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency, *Yale Economic Essays 9*, 3-45.23.

Fama, Eugene F., and Kenneth R. French. 1992, The Cross-Section of Expected Stock Returns, *The Journal of Finance Vol. XLVII*, No.2.

Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics 33*, 3-56.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy 81*, 607-636.

Fu, Fangjian, 2009, Idiosyncratic Risk and the Cross-Section of Expected Stock Returns, *Journal of Financial Economics 91*, 24-37.

Giot, Pierre, 2005, Relationships between Implied Volatility Indexes and Stock Returns, *Journal of Portfolio Management 26*, 12-17.

Goyal, Amit, and Pedro Santa-Clara, 2003, Idiosyncratic Risk Matters! *Journal of Finance 58*, 975-1008.

Goyal, Amit, and Alessio Saretto, 2006, Option Returns and the Cross-Sectional Predictability of Implied Volatility, *working paper, University of Texas at Dallas.*

Lehmann, Bruce N., 1990, Residual Risk Revisited, *Journal of Econometrics 45*, 71-97.

Lintner, John, 1965, The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *Review of Economics and Statistics 47*, 13P37.

Malkiel, Burton G., and Yexiao Xu, 2004, Idiosyncratic Risk and Security Returns, *working paper, University of Texas at Dallas.*

Merton, Robert C., 1987, Presidential Address: A Simple Model of Capital Market Equilibrium with Incomplete Information, *Journal of Finance 42*, 483-510.

Miller, Edward M., 1977, Risk, Uncertainty, and Divergence of Opinion, *Journal of Finance 32*, 1151-1168.

Ofek, Ele, Matthew Richardson, and Robert F. Whitelaw, 2004, Limited Arbitrage and Short Sales Restrictions: Evidence from the Options Markets, *Journal of Financial Economics 74*, 305-342.

Poteshman, Allen M., 2000, Forecasting future volatility from option prices, *working paper, University of Illinois at Urbana-Champaign.*

Ross, Stephen A., 1976, The Arbitrage Theory of Capital Asset Pricing, *Journal of Economic Theory 13*, 341-360.

Sharpe, William F., 1964, Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance 19*, 425-442.

Spiegel, Matthew I. and Xiaotong Wang, 2005, Cross-sectional Variation in Stock Returns: Liquidity and Idiosyncratic Risk, *working paper, Yale University.*

Wei, Steven X., Chu Zhang, 2005, Idiosyncratic Risk Does Not Matter: A Re-examination of the Relationship Between Average Returns and Average Volatilities, *Journal of Banking and Finance 29*, 603-621.

3.8 Tables

Table 3.1: Descriptive Statistics of the Option Firm Sample

This table shows the descriptive statistics for the full sample and three sub-samples in this paper. Panel A is for the full sample period of January 1996 to December 2006 and for stocks with options traded on them on NYSE/AMEX/NASDAQ. There are 3158 firms in total for Panel A. In Panel B, the complete sample is truncated to the period of January 1996 to June 2005. Only NYSE and AMEX stocks are chosen for Panel C and only NYSE stocks are used in Panel D. All ETFs, foreign, financial and utilities stocks are excluded in Panel B to D. Sample mean, median, 5% percentile and 95% percentile are reported. #firms/month is the number of firms in each month and #obs/firm is number of monthly observations per firm. All variables in this table are measured for the one-month period from after option expiration date in month $t-1$ to option expiration date in month t. $SIZE$ is the market value of equity measured in thousand of millions. B/M is Book value-to-Market equity ratio. $Beta$ is the slope in market model. RV is the realized variance of daily underlying stock returns over a month. $ImpVar$ is the implied variance of the one-month-to-expiration options. IV is the total realized idiosyncratic variance which is further decomposed to implied idiosyncratic variance (IIV) and unexpected idiosyncratic variance (UIV).

	Panel A: NYSE/AMEX/NASDAQ, January 1996 to December 2006									
	#firms/month	#obs/firm	$SIZE$	B/M	$Beta$	RV	$ImpVar$	IV	IIV	UIV
Mean	1588	66	4.08	1.06	1.18	27.84%	31.80%	24.95%	24.51%	0.43%
Median	1497	61	0.92	0.40	1.07	18.65%	21.48%	16.53%	16.04%	0.49%
5th Pctl	886	3	0.10	0.08	0.48	2.65%	4.04%	0.97%	1.64%	-15.11%
95th Pctl	2484	131	14.94	1.69	2.16	83.50%	85.94%	75.68%	69.70%	17.48%
	Panel B: NYSE/AMEX/NASDAQ, January 1996 to June 2005									
	#firms/month	#obs/firm	$SIZE$	B/M	$Beta$	RV	$ImpVar$	IV	IIV	UIV
Mean	1257	66	4.63	0.46	1.13	36.11%	39.12%	32.22%	30.05%	2.18%
Median	1267	67	0.90	0.40	0.99	24.57%	28.03%	22.29%	21.24%	1.51%
5th Pctl	794	6	0.16	0.09	0.45	6.66%	7.51%	5.14%	4.48%	-12.00%
95th Pctl	1663	112	16.82	1.01	2.20	104.26%	99.71%	93.02%	79.76%	24.97%
	Panel C: NYSE/AMEX, January 1996 to December 2006									
	#firms/month	#obs/firm	$SIZE$	B/M	$Beta$	RV	$ImpVar$	IV	IIV	UIV
Mean	790	83	6.88	0.48	1.00	18.05%	20.49%	16.00%	15.78%	0.22%
Median	784	93	1.80	0.41	0.92	14.05%	15.96%	12.01%	11.57%	0.56%
5th Pctl	551	7	0.27	0.09	0.46	4.70%	6.08%	3.70%	3.59%	-7.42%
95th Pctl	1057	131	26.75	1.02	1.83	43.13%	51.06%	40.33%	41.36%	8.18%
	Panel D: NYSE only, January 1996 to December 2006									
	#firms/month	#obs/firm	$SIZE$	B/M	$Beta$	RV	$ImpVar$	IV	IIV	UIV
Mean	780	86	7.15	0.49	0.99	16.78%	18.60%	14.71%	14.01%	0.69%
Median	778	97	1.89	0.42	0.91	13.61%	15.47%	11.61%	11.26%	0.63%
5th Pctl	544	8	0.34	0.10	0.47	4.60%	5.99%	3.67%	3.57%	-6.03%
95th Pctl	1027	131	27.10	1.04	1.81	40.29%	42.35%	36.59%	34.36%	7.95%

Table 3.2: Forecasting Future Realized Idiosyncratic Variance

This table shows the tests of the forecasting power of implied idiosyncratic variance (IIV_t).The coefficients for regressions (3.6)-(3.8) in section 3.3.3 are reported and t-statistics are in parentheses. The dependent variable IV_{t+1} is the future realized idiosyncratic variance. IV_t is the past realized idiosyncratic variance. Only stocks with at least 30 observations are included in this analysis. This gives 2180 firms in the final sample.

Dependent Variable: IV_{t+1}			
IIV_t	0.5804**		0.472**
	(67.73)		(52.48)
IV_t		0.3361**	0.1816**
		(76.76)	(40.52)
Adj. R^2	17.64%	14.02%	22.15%

Note: **: significant at 1% level; *: significant at 5% level

Table 3.3: Cross-sectional Regressions of Stock Returns on Idiosyncratic Volatilities

The table reports the coefficient estimates of the Fama and MacBeth (1973) cross-sectional regressions with t-statistics reported in parentheses. The sample is composed of stocks with tradable options on NYSE/AMEX/NASDAQ during the 11-year period of January 1996 through December 2006.The dependent variable is the future one-month holding period return of the stocks . The contemporaneous implied (IIV_t), realized(IV_{t+1}), lagged unexpected (UIV_t) idiosyncratic variances and other lagged variables are used as independent variables to examine the relation between stock return and idiosyncratic volatility. Natural log of SIZE(LnSIZE), natural log of Book-to-market(LnB/M), one-month lagged return (r_t) and Beta from the market model are controlled for. All returns and variances are annualized. Adjusted R-square is reported as Adj.R^2.

	Dependent Variable: r_{t+1}						
X	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IIV_t	-0.0038	-0.0183	0.0101				
	(-0.05)	(-0.25)	(0.14)				
IIV_{t-1}					0.0304		-0.044
					(0.44)		(-0.55)
IV_t			-0.0687*	-0.0733*	-0.0743*	-0.044	
			(-2.27)	(-2.03)	(-2.52)	(-0.55)	
UIV_t		-0.0456				-0.0304	-0.0743*
		(-1.88)				(-0.43)	(-2.52)
LnSIZE	-0.0331**	-0.0335**	-0.0358**	-0.0375**	-0.0345**	-0.0345**	-0.0345**
	(-3.07)	(-3.14)	(-3.45)	(-3.24)	(-3.64)	(-3.64)	(-3.64)
LnB/M	0.0248	0.0246	0.0228	0.0234	0.0237	0.0237	0.0237
	(1.85)	(1.85)	(1.81)	(1.67)	(1.82)	(1.82)	(1.82)
r_t	-0.0258**	-0.0266**	-0.0272**	-0.0267**	-0.0273**	-0.0273**	-0.0273**
	(-4.23)	(-4.32)	(-4.91)	(-4.30)	(-4.38)	(-4.38)	(-4.38)
Beta	0.0309	0.0338	0.0346	0.0331	0.0323	0.0323	0.0323
	(0.86)	(0.94)	(0.99)	(0.96)	(0.96)	(0.96)	(0.96)
Adj. R^2	6.79%	6.95%	7.05%	6.30%	6.93%	6.66%	6.93%

Note: **: significant at 1% level; *: significant at 5% level

Table 3.4: Cross-sectional Regressions of Stock Returns on Idiosyncratic Volatilities (Jan 1996-June 2005)

The table reports the coefficient estimates of the Fama and MacBeth (1973) cross-sectional regressions with t-statistics reported in parentheses. The sample is composed of stocks with tradable options on NYSE/AMEX/NASDAQ during the period of January 1996 through June 2005.All ETFs, foreign, financial and utility stocks are excluded. The dependent variable is the future one-month holding period return of the stocks. The contemporaneous implied (IIV_t), realized(IV_{t+1}), lagged unexpected (UIV_{t+1}) idiosyncratic variances and other lagged variables are used as independent variables to examine the relation between stock return and idiosyncratic volatility. Natural log of SIZE(LnSIZE), natural log of Book-to-market(LnB/M), one-month lagged return r_t and Beta from the market model are controlled for. All returns and variances are annualized. Adjusted R-square is reported as Adj.R^2.

Dependent Variable: r_{t+1}							
X	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IIV_t	0.0573	0.0391	0.0774				
	(0.70)	(0.48)	(0.98)				
IIV_{t-1}					0.0799		-0.0122
					(1.00)		(-0.13)
IV_t			-0.0946**	-0.0844*	-0.0921**	-0.0122	
			(-2.75)	(-2.07)	(-2.80)	(-0.13)	
UIV_t		-0.0524				-0.0799	-0.0921**
		(-1.94)				(-1.00)	(-2.80)
LnSIZE	-0.0318*	-0.0325**	-0.0356**	-0.0412**	-0.0357**	-0.0357**	-0.0357**
	(-2.56)	(-2.65)	(-3.03)	(-2.98)	(-3.17)	(-3.17)	(-3.17)
LnB/M	0.0425	0.0421	0.0394	0.0394	0.0398	0.0398	0.0398
	(1.97)*	(1.97)*	-1.86	-1.73	-1.85	-1.85	-1.85
r_t	-0.0242	-0.0254	-0.0261	-0.026	-0.0267	-0.0267	-0.0267
	(-3.72)**	(-3.86)**	(-3.99)**	(-3.88)**	(-3.97)**	(-3.97)**	(-3.97)**
Beta	0.0314	0.0361	0.0376	0.0351	0.0355	0.0355	0.0355
	-0.77	-0.88	-0.94	-0.9	-0.93	-0.93	-0.93
Adj. R^2	7.72%	7.90%	8.01%	7.30%	8.00%	7.73%	8.00%

Note: **: significant at 1% level; *: significant at 5% level

Table 3.5: Cross-sectional Regressions of Stock Returns on Idiosyncratic Variances (NYSE/AMEX stocks only)

The table reports the coefficient estimates of the Fama and MacBeth (1973) cross-sectional regressions with t-statistics reported in parentheses. The sample is composed of stocks with tradable options on NYSE/AMEX during the period of January 1996 through June 2006.All ETFs, foreign, financial and utility stocks are excluded. The dependent variable is the future one-month holding period return of the stocks. The contemporaneous implied (IIV_t), realized(IV_{t+1}), lagged unexpected (UIV_{t+1}) idiosyncratic variances and other lagged variables are used as independent variables to examine the relation between stock return and idiosyncratic volatility. Natural log of SIZE (LnSIZE), natural log of Book-toBmarket equity (LnB/M), one-month lagged return (r_t) and Beta from the market model are controlled for. All returns and variances are annualized. Adjusted R-square is reported as Adj.R^2.

	Dependent Variable: r_{t+1}						
X	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IIV_t	0.1404	0.1207	0.1143				
	(1.58)	(1.35)	(1.3)				
IIV_{t-1}					0.0722		0.098
					(0.83)		(0.77)
IV_t			-0.005	0.0206	0.0044	0.0765	
			(-0.10)	(0.39)	(0.09)	(0.77)	
UIV_t		-0.0017				-0.0722	-0.0017
		(-0.04)				(-0.83)	(-0.09)
LnSIZE	-0.0272**	-0.0269**	-0.0279**	-0.0324**	-0.0298**	-0.0298**	-0.0298**
	(-3.27)	(-3.25)	(-3.49)	(-3.73)	(-3.80)	(-3.80)	(-3.80)
LnB/M	0.0463**	0.0463**	0.0453**	0.047**	0.0443**	0.0443**	0.0443**
	(3.23)	(3.23)	(3.21)	(3.21)	(3.14)	(3.14)	(3.14)
r_t	-0.0278**	0.0463**	-0.028**	-0.0299**	-0.0306**	-0.0306**	-0.0306**
	(-3.71)	(-3.58)	(-3.77)	(-4.04)	(-4.09)	(-4.09)	(-4.09)
Beta	0.0513	0.0496	0.0498	0.0467	0.0477	0.0477	0.0477
	(1.53)	(1.48)	(1.53)	(1.47)	(1.5)	(1.5)	(1.5)
Adj. R^2	5.78%	5.99%	6.08%	5.46%	6.08%	6.08%	6.08%

Note: **: significant at 1% level; *: significant at 5% level

Table 3.6: Cross-sectional Regressions of Stock Returns on Idiosyncratic Variances (NYSE stocks only)

The table reports the coefficient estimates of the Fama and MacBeth (1973) cross-sectional regressions with t-statistics reported in parentheses. The sample is composed of stocks with tradable options on NYSE during the period of January 1996 through June 2006. All ETFs, foreign, financial and utility stocks are excluded. The dependent variable is the future one-month holding period return of the stocks. The contemporaneous implied (IIV_t), realized(IV_{t+1}), lagged unexpected (UIV_{t+1}) idiosyncratic variances and other lagged variables are used as independent variables to examine the relation between stock return and idiosyncratic volatility. Natural log of SIZE (LnSIZE), natural log of Book-to-market equity (LnB/M), one-month lagged return (r_t) and Beta from the market model are controlled for. All returns and variances are annualized.Adjusted R-square is reported as Adj.R^2.

Dependent Variable: r_{t+1}							
X	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IIV_t	0.2410*	0.229*	0.2124*				
	(2.52)	(2.36)	(2.20)				
IIV_{t-1}					0.145		0.1737
					(1.58)		(1.71)
IV_t			0.0131	0.0642	0.0287	0.1737	
			(0.25)	(1.19)	(0.56)	(1.71)	
UIV_t		-0.0109				-0.145	0.0287
		(-0.24)				(-1.58)	(0.56)
Ln(SIZE)	-0.0277**	-0.0276*	-0.0283*	-0.0353*	-0.0309*	-0.0309*	-0.0309*
	(-3.33)	(-3.34)	(-3.53)	(-4.08)	(-3.93)	(-3.93)	(-3.93)
Ln(B/M)	0.0431**	0.0429**	0.0418**	0.0433**	0.0415**	0.0415**	0.0415**
	(3.04)	(3.05)	(3.01)	(3.02)	(2.98)	(2.99)	(2.99)
r_t	-0.0277**	-0.0266*	-0.0278**	-0.0299**	-0.031**	-0.031**	-0.031**
	(-3.74)	(-3.59)	(-3.77)	(-4.04)	(-4.15)	(-4.15)	(-4.15)
Beta	0.0537	0.0517	0.0498	0.046	0.0468	0.0468	0.0468
	(1.58)	(1.53)	(1.51)	(1.43)	(1.46)	(1.46)	(1.46)
Adj. R^2	5.80%	6.02%	6.09%	5.43%	6.06%	6.06%	6.06%

Note: **: significant at 1% level; *: significant at 5% level

Table 3.7: Cross-sectional Regressions of Stock Return on Persistent UIV

This table present coefficient estimates and t-statistics (in parentheses) for Fama-MacBeth (1973) regressions. The dependent variable is the future stock return r_{t+1}. The independent variables are current implied idiosyncratic variance (IIV_t), past realized idiosyncratic variance (IV_{t-1}) and the past unexpected idiosyncratic variance (UIV_{t-1}). Natural log of SIZE (LnSIZE), natural log of Book-to-market equity (LnB/M), one-month lagged return (r_t) and Beta from the market model are controlled for. All returns and variances are annualized. Adjusted R-square is reported as Adj.R^2.

	Dependent Variable: r_{t+1}			
X	Complete Sample	DDP Sample	NYSE/AMEX Sample	NYSE Sample
IIV_t	-0.0286	0.0324	0.1078	0.1984*
	(-0.40)	(0.41)	(1.17)	(1.98)
IV_{t-1}	-0.0455	-0.0657	-0.0157	0.0005
	(-1.48)	(-1.84)	(-0.32)	(0.01)
UIV_{t-1}	-0.0496	-0.0596*	-0.004	-0.0421
	(-1.90)	(-2.03)	(-0.10)	(-1.00)
LnSIZE	-0.0345**	-0.034**	-0.0267**	-0.0268**
	(-3.41)	(-2.94)	(-3.38)	(-3.40)
LnB/M	0.0231	0.0409	0.0441**	0.0409**
	(1.76)	(1.92)	(3.10)	(2.93)
r_t	-0.0278**	-0.0271**	-0.0282**	-0.0281**
	(-4.52)	(-4.08)	(-3.37)	(-3.76)
Beta	0.034	0.0376	0.0512	0.0527
	(0.97)	(0.94)	(1.55)	(1.58)
Adj. R^2	7.20%	8.19%	6.32%	6.31%

Note: **: significant at 1% level; *: significant at 5% level

Appendix A

Extension of the Multi-Stage Model

Extension of the Multi-Stage Model

This appendix presents an extension of the multi-stage model with an initial debt coupon. Let C_0 denote the initial coupon, and Π_0 the scale of the firm in stage 0. As the firm has debt in place in stage 0, there is also an endogenous default threshold is X_{D0}. The values of this initial debt issue in stages 0, 1, 2 are denoted by D_0, D_{10}, and D_{20}. Other variables are the same as in Section 2.2.

Mature Firm (Stage 2)

In the second stage, the investment options have been exercised, so the firm faces a pure financing decision. The new debt D_{22} is issued in this stage to partially finance the investment cost F_2 and equityholders bear the remainder of the cost. The new debt D_{22} together with the debt issued in the first stage D_{21} and the initial debt D_{20} offers tax savings but creates default risk. The solutions of the debt values are simple generalizations of equation (2.6). We again assume that D_{22} has the same seniority as D_{21} and D_{20}. The values of the three debt issues for $X \geq X_{D2}$ are given by:

$$\begin{aligned} D_{2i}(X, C_0, C_1, C_2) &= \frac{C_i}{r}\Big(1 - (\frac{X}{X_{D2}})^a\Big) + \\ &\quad \frac{C_i}{C_0 + C_1 + C_2}\frac{(1-\alpha)(1-\tau)(\Pi_0 + \Pi_1 + \Pi_2)X_{D2}}{r-\mu}(\frac{X}{X_{D2}})^a, \end{aligned} \tag{A.1}$$

where $i = 0, 1, 2$. The value of equity for $X \geq X_{D2}$ can be obtained similarly:

$$\begin{aligned} E_2(X, C_0, C_1, C_2) &= (1-\tau)\Big(\frac{(\Pi_0 + \Pi_1 + \Pi_2)X}{r-\mu} - \frac{C_0 + C_1 + C_2}{r} - \\ &\quad (\frac{(\Pi_0 + \Pi_1 + \Pi_2)X_{D2}}{r-\mu} - \frac{C_0 + C_1 + C_2}{r})(\frac{X}{X_{D2}})^a\Big). \end{aligned} \tag{A.2}$$

The only decision that the firm's equityholders make in stage 2 is when to default. To maximize the value of this option, equityholders select an endogenous default threshold X_{D2} such that:

$$\frac{\partial E_2(X, C_0, C_1, C_2)}{\partial X}\Big|_{X = X_{D2}} = 0, \tag{A.3}$$

which yields a closed-form solution for the optimal default threshold in the second stage:

$$X_{D2} = \frac{a(C_0 + C_1 + C_2)(r-\mu)}{r(a-1)(\Pi_0 + \Pi_1 + \Pi_2)}. \tag{A.4}$$

Adolescent Firm (Stage 1)

In the first stage, the first investment option has been exercised. The adolescent firm has some assets-in-place and its capital structure is a mix of debt, D_{10} and D_{11}, and equity, E_1. It has both an option to default and an option to invest, so it solves a joint financing and investment problem.

Using similar arguments as in Section 2.2, each of the valuation equations for D_{10}, D_{11} and E_1 needs to satisfy two boundary conditions. Consider debt $D_{1i}(X, C_0, C_1, C_2)$, with $i = 0, 1$. When $X \downarrow X_{D1}$, equityholders default and debtholders get the liquidation value:

$$D_{1i}(X_{D1}, C_0, C_1, C_2) = \frac{C_i}{C_0 + C_1} \frac{(1-\alpha)(1-\tau)(\Pi_0 + \Pi_1)X_{D1}}{r - \mu}. \tag{A.5}$$

If the firm keeps growing and X increases to the investment threshold X_{S2}, the firm will exercise the second-stage investment option, and debt values from stage 1 equal debt values in stage 2: $D_{1i}(X_{S2}, C_0, C_1, C_2) = D_{2i}(X_{S2}, C_0, C_1, C_2)$. For $X_{D1} \leq X \leq X_{S2}$, these conditions imply the following solution for debt value in stage 1:

$$\begin{aligned} D_{1i}(X, C_0, C_1, C_2) \;=\; & \frac{C_i}{r}\Big(1 - L(X) - (\frac{X_{S2}}{X_{D2}})^a H(X)\Big) + (1-\alpha)(1-\tau) \\ & \Big(\frac{C_i}{C_0 + C_1}\frac{(\Pi_0 + \Pi_1)X_{D1}}{r - \mu} L(X) + \\ & \frac{C_i}{C_0 + C_1 + C_2}\frac{(\Pi_0 + \Pi_1 + \Pi_2)X_{D2}}{r - \mu}(\frac{X_{S2}}{X_{D2}})^a H(X)\Big), \end{aligned} \tag{A.6}$$

where $i = 0, 1$ and where

$$L(X) = \frac{X^z X_{S2}^a - X^a X_{S2}^z}{X_{D1}^z X_{S2}^a - X_{D1}^a X_{S2}^z} \tag{A.7}$$

$$H(X) = \frac{X_{D1}^z X^a - X_{D1}^a X^z}{X_{D1}^z X_{S2}^a - X_{D1}^a X_{S2}^z} \tag{A.8}$$

denote state prices that, respectively, take the value of one if X first reaches the default threshold X_{D1} from above or the investment threshold X_{S2} from below.

The value of equity, E_1, on the other hand, approaches zero when $X \downarrow X_{D1}$. When $X \uparrow X_{S2}$, it satisfies the value-matching condition $E_1(X_{S2}, C_0, C_1, C_2) = E_2(X_{S2}, C_0, C_1, C_2) - [F_2 - D_{22}(X_{S2}, C_0, C_1, C_2)]$ because the fixed investment cost, F_2, is funded by a mix of debt and equity. For $X_{D1} \leq X \leq X_{S2}$, these conditions imply the following solution for equity value in stage 1:

$$\begin{aligned} E_1(X, C_0, C_1, C_2) \;=\; & (1-\tau)\Big[(\frac{(\Pi_0 + \Pi_1)X}{r - \mu} - \frac{(C_0 + C_1)}{r}) - (\frac{(\Pi_0 + \Pi_1)X_{D1}}{r - \mu} - \frac{C_0 + C_1}{r})L(X) + \\ & \Big(\frac{\Pi_2 X_{S2}}{r - \mu} - \frac{C_2}{r} - \frac{F_2 - D_{22}(X_{S2}, C_0, C_1, C_2)}{1 - \tau} - \\ & (\frac{(\Pi_0 + \Pi_1 + \Pi_2)X_{D2}}{r - \mu} - \frac{C_0 + C_1 + C_2}{r})(\frac{X_{S2}}{X_{D2}})^a\Big)H(X)\Big]. \end{aligned} \tag{A.9}$$

The first two terms in equation (A.9) denote the present value of after-tax cash flows to equityholders until the firm defaults in the current stage. The next few terms in this equation show the value from entering into the second stage. Given E_1, equityholders can determine the optimal default threshold, X_{D1}, by maximizing equity value:

$$\frac{\partial E_1(X, C_0, C_1, C_2)}{\partial X}|_{X=X_{D1}} = 0. \tag{A.10}$$

Furthermore, the optimal investment threshold, X_{S2}, solves the smooth-pasting condition:

$$\frac{\partial E_1(X, C_0, C_1, C_2)}{\partial X}|_{X=X_{S2}} = \frac{\partial E_2(X, C_0, C_1, C_2)}{\partial X}|_{X=X_{S2}} + \frac{\partial D_{22}(X, C_0, C_1, C_2)}{\partial X}|_{X=X_{S2}}. \tag{A.11}$$

Juvenile Firm (Stage 0)

In the initial stage, the juvenile firm now has some assets-in-place, an option on a two-stage investment project, and pre-existing debt. The firm thus faces a joint financing and investment problem. As $X \downarrow X_{D0}$, equityholders default and end up with nothing, $E_0(X_{D0}, C_0, C_1, C_2) = 0$, whereas debtholders receive the liquidation value $D_0(X_{D0}, C_0, C_1, C_2) = (1-\tau)(1-\alpha)\frac{\Pi_0 X_{D0}}{r-\mu}$. When X touches the first investment threshold X_{S1} the first time from below, the first option is exercised and hence:

$$E_0(X_{S1}, C_0, C_1, C_2) = E_1(X_{S1}, C_0, C_1, C_2) - [F_1 - D_{11}(X_{S1}, C_0, C_1, C_2)], \tag{A.12}$$

because debt and equity finance the exercise cost F_1. In addition, the initial debt satisfies the value-matching condition:

$$D_0(X_{S1}, C_0, C_1, C_2) = D_{10}(X_{S1}, C_0, C_1, C_2). \tag{A.13}$$

For $X_{D0} \leq X \leq X_{S1}$, these conditions yield the following solutions for debt and equity values:

$$\begin{aligned} D_0(X, C_0, C_1, C_2) = {} & \frac{C_0}{r}\Big(1 - \tilde{L}(X) - [L(X_{S1}) + (\frac{X_{S2}}{X_{D2}})^a H(X_{S1})]\tilde{H}(X)\Big) + \\ & (1-\alpha)(1-\tau)\Big(\frac{\Pi_0 X_{D0}}{r-\mu}\tilde{L}(X) + \frac{C_0}{C_0+C_1}\frac{(\Pi_0+\Pi_1)X_{D1}}{r-\mu}L(X_{S1})\tilde{H}(X) + \\ & \frac{C_0}{C_0+C_1+C_2}\frac{(\Pi_0+\Pi_1+\Pi_2)X_{D2}}{r-\mu}(\frac{X_{S2}}{X_{D2}})^a H(X_{S1})\tilde{H}(X)\Big), \end{aligned} \tag{A.14}$$

and

$$\begin{aligned} E_0(X, C_0, C_1, C_2) = {} & (1-\tau)\Big[(\frac{\Pi_0 X}{r-\mu} - \frac{C_0}{r}) - (\frac{\Pi_0 X_{D0}}{r-\mu} - \frac{C_0}{r})\tilde{L}(X) + \Big((\frac{\Pi_1 X_{S1}}{r-\mu} - \frac{C_1}{r}) - \\ & \frac{F_1 - D_{11}(X_{S1}, C_0, C_1, C_2)}{1-\tau} - (\frac{(\Pi_0+\Pi_1)X_{D1}}{r-\mu} - \frac{C_0+C_1}{r})L(X) + \\ & (\frac{\Pi_2 X_{S2}}{r-\mu} - \frac{F_2 - D_{22}(X_{S2}, C_0, C_1, C_2)}{1-\tau} - \frac{C_2}{r} - \\ & (\frac{(\Pi_0+\Pi_1+\Pi_2)X_{D2}}{r-\mu} - \frac{C_0+C_1+C_2}{r})(\frac{X_{S2}}{X_{D2}})^a)H(X)\Big)\tilde{H}(X), \end{aligned} \tag{A.15}$$

where $H(X)$ and $L(X)$ are defined in equation (A.7), and where

$$\tilde{L}(X) = \frac{X^z X_{S2}^a - X^a X_{S1}^z}{X_{D0}^z X_{S1}^a - X_{D0}^a X_{S1}^z} \tag{A.16}$$

$$\tilde{H}(X) = \frac{X_{D0}^z X^a - X_{D0}^a X^z}{X_{D0}^z X_{S1}^a - X_{D0}^a X_{S1}^z} \tag{A.17}$$

denote state prices that, respectively, take the value of one if X first reaches the default threshold X_{D0} from above or the investment threshold X_{S1} from below.

Finally, the firm's equityholders will choose an optimal pair (C_1, C_2) by maximizing initial firm value subject to the smooth-pasting conditions for X_{D0}, X_{D1}, X_{D2}, X_{S1} and X_{S2} mentioned above:

$$\max_{C_1,C_2} \; D_0(X_0, C_0, C_1, C_2) \; + \; E_0(X_0, C_0, C_1, C_2) \tag{A.18}$$

subject to:

$$\frac{\partial E_0(X, C_0, C_1, C_2)}{\partial X}\Big|_{X=X_{D0}} \; = \; 0, \tag{A.19}$$

$$\frac{\partial E_1(X, C_0, C_1, C_2)}{\partial X}\Big|_{X=X_{D1}} \; = \; 0, \tag{A.20}$$

$$\frac{\partial E_2(X, C_0, C_1, C_2)}{\partial X}\Big|_{X=X_{D2}} \; = \; 0, \tag{A.21}$$

$$\frac{\partial E_0(X, C_0, C_1, C_2)}{\partial X}\Big|_{X=X_{S1}} \; = \; \frac{E_1(X, C_0, C_1, C_2)}{X}\Big|_{X=X_{S1}} + \frac{D_{11}(X, C_0, C_1, C_2)}{X}\Big|_{X=X_{S1}}, \tag{A.22}$$

$$\frac{\partial E_1(X, C_0, C_1, C_2)}{\partial X}\Big|_{X=X_{S2}} \; = \; \frac{E_2(X, C_0, C_1, C_2)}{X}\Big|_{X=X_{S2}} + \frac{D_{22}(X, C_0, C_1, C_2)}{X}\Big|_{X=X_{S2}}. \tag{A.23}$$

Appendix B

Simulated Method of Moments

Simulated Method of Moments

We calibrate the structural parameters of the model via Simulated Method of Moments (SMM), which is based on indirect inference techniques in Gourieroux, Monfort, and Renault (1993) and Gourieroux and Monfort (1996). By varying the vector of model parameters, b, SMM minimizes the distance between model moments, denoted as $M_m(b)$, and data moments, denoted as M_d. Note that we explicitly state the dependence of the simulated moments, $M_m(b)$, on the vector of structural parameter values, b.

The simulated moments parameter estimation procedure can be summarized as follows (see, e.g., Hennessy and Whited (2005, 2007) for further details):

1. We first compute N_d data moments from COMPUSTAT to generate the vector of data moments, M_d. We use fixed firm and year effects in the estimation of all of our data moments to remove heterogeneity from the actual data.

2. The variance-covariance matrix of the data moments yields the optimal weighting matrix:

$$W_d \;=\; \left[N_d \mathrm{Var}(M_d)\right]^{-1}. \tag{B.1}$$

   Intuitively, this weighing matrix places more weight on more precisely measured moments.

3. For each vector of structural parameter values, b, we simulate a set of S panel data sets with i.i.d. firms each containing $2 * T$ firm-year observations. We discard the first T years of data to avoid non-stationarity and other problems arising from the initial conditions of the simulations. We then calculate the same set of moments as in step 1 using our S simulated panel data sets to generate $M_m(b)$.

4. We then calculate the weighted distance between the model moments and the data moments:

$$J_{N_d}(b) \;=\; \Big[M_d - \frac{1}{S}\sum_{i=1}^{S} M_m(b)\Big]' W_d \Big[M_d - \frac{1}{S}\sum_{i=1}^{S} M_m(b)\Big], \tag{B.2}$$

   where W_d is the weighting matrix from step 2.

5. Finally, by varying b iteratively, we find an optimal set of structural parameter values, b^*, which minimizes the objective function, $J_{N_d}(b)$:

$$b^* \;=\; \arg\min_b \Big[M_d - \frac{1}{S}\sum_{i=1}^{S} M_m(b)\Big]' W_d \Big[M_d - \frac{1}{S}\sum_{i=1}^{S} M_m(b)\Big]. \tag{B.3}$$