UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Essays in technology adoption and corporate finance

Permalink
https://escholarship.org/uc/item/4s52z1kj

Author
Patel, Pratish

Publication Date
2013

Peer reviewed|Thesis/dissertation

Essays in technology adoption and corporate finance

by

Pratish Anilkumar Patel

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

William Fuchs, Co-chair
Nancy Wallace, Co-chair
Robert Helsley
Benjamin Handel

Spring 2013

Essays in technology adoption and corporate finance

Copyright 2013

by

Pratish Anilkumar Patel

Abstract

Essays in technology adoption and corporate finance

by

Pratish Anilkumar Patel

Doctor of Philosophy in Business Administration

University of California, Berkeley

William Fuchs, Co-chair

Nancy Wallace, Co-chair

This dissertation consists of three chapters that concern technology adoption and corporate finance. The first chapter analyzes the optimal investment strategy of two firms confronted with the option to adopt a new technology. I add two key features: location and learning. A firm gains relative advantage entirely due to its geographic placement — this is the location benefit. Firms also learn from the adoption experience of their rival — this is the learning benefit. Imperfect competition induces firms to adopt early while learning induces firms to wait. This tradeoff has two implications: First, firms in better locations should never adopt after their rivals; second, technology adoption should be geographically clustered. These implications are consistent with the direct evidence regarding technology adoption. Since investment in a new technology is a growth option, location and learning also affect asset prices. I show that firms risk loadings (βs) and returns correlate positively for geographically close firms.

The second chapter analyzes the link between debt maturity and term spread. This chapter is co-authored with Paulo Issler. Evidence shows that firm's debt maturity and term spread are intricately linked. Firms issue short term debt when the term spread is significantly positive and they increase maturity as the term spread decreases. The current literature explains this link with market frictions such as agency problems, asymmetric information, and liquidity risk. We explain the link between debt maturity and term spread using the trade-off theory

of capital structure. When the term spread is small or even negative, transaction costs of debt rollover outweigh bankruptcy costs. Therefore, the firm optimally chooses to increase debt maturity. On the other hand, when the term spread is significantly positive, bankruptcy costs outweigh transaction costs of debt rollover. Therefore shorter debt maturity is optimal as it minimizes the chance of bankruptcy. In addition, we contribute to the current discussion in the literature concerning the speed of adjustments of capital structure. Our results show that firms are active in adjusting their capital structure. The model is consistent with a variety of stylized facts concerning debt maturity.

The third chapter analyzes the role of debt financing on skyscraper heights. This chapter is co-authored with Robert Helsley. Skyscraper clusters, which form a city's skyline, give a city its unique identity. Despite their ubiquitous nature, theory is scarce. The traditional model in urban economics attributes the existence of tall buildings to agglomeration. However, the model ignores two important economic forces: debt financing and imperfectly competitive real estate markets. In this article, we develop a game-theoretic model that captures these two forces. A skyscraper owner builds taller because she enjoys limited liability. We motivate the model with a case study. We analyze John Raskob's letter (owner of the Empire-State Building) to his financier Louis Kauffman; this letter details the underlying role of debt financing behind the height choice. Debt financing along with imperfect competition make the skyscraper owner pursue an aggressive strategy. The owner, who maximizes equity, builds much taller than what is optimal in autarky without debt financing.

Dedication

I dedicate my Ph.D. thesis to two people who are not present at this moment. First, I would like to dedicate my work to the memory of my wonderful cousin: Amaey Shah. He was nine years old when he passed away from cancer. His enthusiasm for life and the pursuit of knowledge will always remain with me. Specifically, I am reminded of our last conversation. I asked him about the most favorite food of vegetarian dinosaurs. We both talked about some obvious choices but could not decide on the absolute favorite. Afterward, we laughed off the question. Now, Amaey knows the answer to that question.

Second, I would like to dedicate my work to my unborn child nicknamed Bunzers. I recently found out that I am going to be a father. The thought of me being a father really excites me. I cannot wait to teach my kid about "real options" and the associated optimal stopping problem on a hike somewhere.

Contents

List of Figures

List of Tables

Acknowledgments

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me.

Above all, I would like to thank my wife Bharvi for her personal support and great patience at all times.

I also thank my parents Anil and Harsha, my sister Ruchi, my parents in law Bharat and Prabha, my uncle Apurva, my aunt Purvi, and my cousin Arjun for their love, support, and encouragement throughout my life.

I would have especially like to thank Paulo Issler, Matteo Maggiori, Raymond Leung, Brian Ayash, Nirupama Kulkarni, Sean Wilkoff, Andres Donangelo, Barney Glaesar, Boris Albul, Narhari Phatak, Issi Romem, Jiakai Chen, Michael Weber, Isaac Hacamo, and Donatella Taurasi. You have made my Ph.D. tenure at Berkeley an absolute joy.

I am deeply grateful to my advisor William Fuchs. I know that my job market paper is not part of mainstream finance literature. Yet, at no point in time, you discouraged me from pursuing my idea. You have listened to my ideas and discussions with you frequently led to key insights. Your ability to convey a complex idea in a simple manner is something I will take with me going forward.

I am also grateful to my co-advisor Nancy Wallace. You have always been supportive of my research ideas. In the third year, I was interested in theoretical econometrics and you were completely on board with me. With my job market paper, you were extremely influential in framing the question and pushing me to make the paper more appealing to finance audience. Above all, you have always given me grant support and even with your extremely busy schedule, you have always made time for me. I deeply appreciate the countless hours you put in while I was going through my job market. During my Ph.D. career, I realized that I was poor at judging "a good paper". You have shown me some steps on how to improve my bad taste. Above all, your passion for real estate related issues is something that I really admire — I hope to incorporate some of this in my career.

I am also grateful to my committee member Robert Helsley. You introduced me to the beautiful field of urban economics — Thank you. You had once told me something to the effect of "We are going to be remembered by not our work but by our writing". At that moment, I did not think much of your comment, but now I have a newfound respect for good writers. As a result, one of my hobbies is to highlight a "good sentence". Furthermore, I really enjoyed our sessions in which you let me work out the math on your white board. I am a theorist at heart and you taught me how to write a good theory paper. Thank you so

Chapter 1

The Adoption of New Technologies: Location, Learning and Asset Pricing Implications

1.1   Introduction

One of the most striking features of the geography of economic activity is its concentration—production is remarkably concentrated in space Krugman (1991). There is a small but burgeoning literature showing that technology adoption is also geographically concentrated (Kelley and Helper (1999), Baptista (2000) and No (2008)). Firms and workers are much more productive in large and dense urban environments than other locations. Both Smith (1904) and Marshall (1961) recognized the value of such dense urban environments that allow workers to interact more frequently. Workers learn from each other, and in the process of exchanging ideas, they effectively transfer knowledge regarding new technology. In this paper, I investigate the implications of location and learning on technology adoption by analyzing the geographical patterns of technology adoption and asset prices.

In the model, two firms have the option to adopt a new technology that reduces marginal cost of production. There are three key ingredients: competition, location and learning. First, consider the effect of competition. Adoption increases a firm's flow of earnings, but adopting sooner is more expensive than adopting later. Under imperfect competition, since the earnings of a firm depend on its rivals, both firms engage in a value-destroying strategy of early adoption. Next, consider the effect of location. A firm's location of production gives it a relative advantage. For example, one firm may be closer to an urban cluster and hence have a higher access to a pool of skilled workers. Traditionally, such location benefits are motivated with the label "transportation cost". I assume that firm 1—which is in a better location—has a lower transportation cost than firm 2. This is the only source of asymmetry

between firms.

Finally, consider the effect of learning. The cost of new technology has two components. The first component is the actual physical cost of technology. The second component is the *soft information* associated with the effective operation of the new technology. By definition, it is difficult to document this soft information and hence it can only be effectively transferred through interpersonal interactions.[1] The importance of the transfer of such soft information should not be underestimated. In a study of the adoption of machinery equipment technology, Teece (1977) estimates the cost of such transfers to be on average 20% of the total project costs. Learning decreases adoption cost.

Frequent interpersonal interactions are plausible when workers are embedded in a network that may be due to either geographical or social ties. It is natural to assume that the frequency of interactions between firms' workers is high when the firms are near each other. Similarly, the frequency of interactions is high when the workers share common ties; for example, if they are alumni of the same university. Lastly, note that learning has a temporal dimension. Workers learn from prior adoption experience of their rival, which can only take place if firms *sequentially* invest; learning is impossible if both firms *simultaneously* invest.

Competition, location and learning cause the following trade-off: On one hand, early investment in the new technology by a firm leads to higher short term earnings. A portion of the increase in earnings of the leader, i.e., the adopting firm, comes at the expense of the follower. Due to its better location, firm 1 has a higher market share and hence is more likely to be the leader because it has a greater ability to appropriate the benefits from adoption. On the other hand, since the follower firm learns from the adoption experience of the leader, it benefits from the lower adoption cost. Learning increases the likelihood that the follower will adopt in the future, which in turn reduces the present value of the leader's earnings.

The tradeoff leads to two pure strategy equilibria: one featuring sequential investment (SEQ) and the other featuring simultaneous investment (SIM). First, when firms are far from each other, investment in the new technology is sequential. Firm 1 invests first and after a significant lag, firm 2 follows suit. Observationally, technology adoption is geographically dispersed. Technology trickles down from firms located in better locations to others less well situated. Second, when firms are near each other, both firms simultaneous invest in the new technology. Observationally, technology adoption is geographically clustered; the lag in adoption timing is zero.

The intuition behind "it is easier to collude among equals" captures the essence of the paper. First, consider the effect of competition and location while ignoring learning. When firms are near each other, the market share of firm 1 is not too different from firm 2. Therefore, firms

---

[1]Note that this information is not the same as the imperfect information as in Grenadier (1999).

tacitly collude and they decide to simultaneously invest. When firms are geographically far from each other, the market share of firm 1 is much higher than firm 2. Colloquially, firm 1 is a "Maverick"—it is unwilling to participate in any collusive action. Now add learning. Since firm 2 learns from firm 1's adoption experience, firm 1 is more prone to collude and simultaneously invest.[2]

Learning plays both a direct and an indirect role. In SEQ equilibrium, firm 2 learns from the adoption experience of firm 1—this is the direct role. When firms are geographically close, frequency of interactions between workers is high and hence learning benefits are high. But when firms are geographically close, both firms simultaneously invest and in this outcome there is no learning. This seemingly counterintuitive result arises because learning also serves as a threat—this is the indirect role. Firm 1 is afraid that it will not be the sole leader for long when learning benefits are high and hence it will not adopt early.

The decision to adopt a new technology is ultimately a growth option possessed by both firms. Competition, location and learning affect the investment strategy of both firms and therefore they also have asset pricing implications. First, consider the case in which firms are geographically distant and when firm 1 has already adopted. Firm 2 has the growth option to adopt the new technology. Any good news in the product market has an asymmetric effect on both firms. Positive news increases the probability that firm 2 will adopt and hence it decreases the future earnings of firm 1. Therefore, competition acts like a natural hedge. Second, consider the case in which firms are geographically near each other. Both firms simultaneously invest in the new technology—the strike price of the option is the same. Therefore, any positive news affects both firms symmetrically. This leads to two testable implications that highlight the impact of geography on asset pricing. When firms are geographically close, both firm risk (β) and stock returns correlate positively.

To summarize, the model is consistent with a variety of stylized facts regarding technology adoption and asset pricing. First, learning from interactions positively affects the probability of technology adoption, consistent with the evidence summarized in Young (2009). Second, the model predicts positive correlation between size and the probability of technology adoption, consistent with Karshenas and Stoneman (1993), Stoneman and Kwon (1996), Baptista (2000), and Hall and Khan (2003). Third, the model predicts geographical clustering of technology adoption, consistent with Kelley and Helper (1999), Baptista (2000), and No (2008). Fourth, the model relies on the fact that adopting first is more beneficial than adopting second, which is consistent with Karshenas and Stoneman (1993), Stoneman and Kwon (1996), and Baptista (2000). Fifth, the model generates positive correlation of stock returns among geographically close firms, consistent with Pirinsky and Wang (2006), Eckel et al. (2011), Barker and Loughran (2007), and Wongchoti and Wu (2008). Finally, the model predicts

---

[2]I use the terminology of "Maverick" from Ivaldi et al. (2003) who present a report concerning the economics of tacit collusion to the European Commission.

that stock returns co-move together in less concentrated industries, consistent with Hoberg and Phillips (2010), and Bustamante (2011).

The paper is organized as follows: Section I summarizes previous work on technology adoption and its asset pricing implications. Section II introduces a simple two-period deterministic model, which provides the basic intuition of the paper. Section III derives equilibrium and considers its properties. Section IV summarizes major empirical implications of the model and relates them to existing evidence. In Section V, I generalize the two period setup by introducing an infinite period stochastic model, which allows me to derive asset pricing implications. Section VI concludes the paper.

1.2   Related Literature

Ever since the pioneering studies of hybrid corn adoption by Ryan and Gross (1943) and Griliches (1957), the empirical literature in economics has emphasized the impact of learning from interpersonal interactions on technology adoption. Ryan and Gross (1943) found that the most important determinant of adoption of hybrid seeds by Iowa farmers is their interactions with neighbors. The importance of learning from interactions has since been corroborated in various studies of technology adoption across industries and across countries. Young (2009) provides a survey of the empirical findings.

The importance of these interactions has also been documented in an international setting. Economic activities such as trade and foreign direct investment are positively correlated with technology adoption.[3] Keller (2002), Abreu, Groot, and Florax (2004), Comin and Hobijn (2004), Comin and Hobijn (2010) and Comin, Dmitriev, and Rossi-Hansberg (2012) find that technology adoption trickles down geographically. Firms located in countries close to adoption leaders adopt first and then the technology slowly trickles down to firms located in countries farther away.

The intuition behind the effect of learning from interactions is straightforward, but in isolation it ignores competition. In a study of the adoption of computerized numerically controlled machines, Stoneman and Kwon (1996) finds that profits of non-adopters decrease as their rivals adopt. There is overwhelming evidence that firm size and the probability of adoption are positively correlated (Stoneman (2002)). Grenadier (1996) illustrates how imperfect competition can lead to boom and bust cycles in real-estate. The seminal paper that incorporates the effect of competition and technology adoption is Fudenberg and Tirole (1985) (hereafter referred to as FT). In their model, FT develop the adoption decision of two firms in an infinite period deterministic setup. In a continuous time setting, it is not obvious what it means when a rival instantaneously reacts to the competitor's actions. FT formally solve

---

[3]Caselli and Coleman (2001) documents positive correlation between trade openness and computer adoption across countries

this issue and derive a *closed-loop equilibrium*—a sub-game perfect equilibrium where firms do not pre-commit to their strategies. Smit and Ankum (1993) and Kulatilaka and Perotti (1998) use a similar framework with imperfect competition and analyze the equilibrium strategies with uncertainty. They were one of the first to show that the option value to wait decreases with competition. Chevalier-Roignant et al. (2011) provides an excellent survey of the literature dealing with competition and investment.

Carlson et al. (2011) and Bustamante (2011) extend FT's model by deriving asset pricing implications.[4] They both clarify the role of imperfect competition on asset prices. For example, Bustamante (2011) (along with Hoberg and Phillips (2010)) find that stock returns in less concentrated industries co-move together. Bena and Garlappi (2010) find an analogous result in a similar setup that links technological innovation and asset prices.

I add location and learning to FT's framework. This allows me to incorporate the evidence regarding both learning and competition on technology adoption. Additionally, I also highlight how geography impacts asset pricing.

1.3 A deterministic model of technology adoption

This section develops a deterministic two-period model to analyze the adoption decision of each of the two firms. I augment FT's framework by adding the effect of location and learning from interpersonal interactions. I highlight the positive effect of market size or market share on technology adoption by explicitly considering product market competition.

The setup

Exogenous demand in the product market

Consider an environment with two periods, $t \in \{0, T\}$, where T is the length of time spanned by the two periods. Firms 1 and 2 are rivals in the homogenous product market where they produce quantities q_1 and q_2 in each period t.

For tractability, I assume that the demand curve is linear. Denoting the market clearing price in period t by P, the inverse demand function is

$$P(X_t, Q_t) = \underline{X} + X_t - Q_t,$$

---

[4]Numerous studies develop the link between asset pricing and real options. Starting with the seminal article Berk, Green, and Naik (1999), Gomes, Kogan, and Zhang (2003), Kogan (2004), Zhang (2005), Cooper (2006), Aguerrevere (2009), and Novy-Marx (2011) link real options, product market competition, and asset pricing. These papers assume pure competition or monopoly competition, or singular investment costs, which are slightly different than my setting.

where $Q_t \equiv q_{1t} + q_{2t}$ is the industry output in period t and parameters $X_t \geq 0$ along with $\underline{X}$ cause parallel shifts in the demand level. Parameter $\underline{X}$, whose lower bound is given below, ensures that the demand level is sufficiently high so that both firms produce positive quantities. The demand level in the second period, X_T, is proportional to the first period demand level X_0, so that $X_T = X_0 e^{\mu T}$, where μ is the demand growth rate.

Lastly, management of both firms discount the cash flows in period T by the discount factor e^{-rT}. I further assume a transversality like condition: $r > \mu$.

Variable costs

Variable costs are composed of two components: manufacturing cost denoted by m and transportation cost denoted by l_i. Transportation cost arises due to firm's location. I assume that transportation cost of firm 1, which is in a better location, is lower than that of firm 2, so that $l_1 < l_2$. Differences in location is the only asymmetry between the two firms.

Initially, both firms use the old technology; they have the same manufacturing costs m. Manufacturing cost of the adopting firm decreases to $m - \alpha$ for $\alpha > 0$. No further technological advances are anticipated.[5]

Effect of learning on the cost of adoption

In order to understand the effect of learning from interpersonal interactions, a distinction must be made in the two components comprising adoption costs.[6]

The first component of the adoption cost concerns the physical equipment itself. For example, consider the adoption of computer numerically controlled (CNC) machines.[7] A CNC machine does not need to be controlled by an operator; it can be programmed to be run by a computer, thereby increasing productivity. The first component is the actual physical cost of CNC machine.

The second component concerns soft information needed to effectively operate the technology. This information consists of the methods of organization and operation, quality control, and other manufacturing procedures. By definition, this soft information is difficult to codify— this information is transferred effectively only through interpersonal interactions. In the example of CNC machines, the second component of the adoption cost is related to the implementation of effective management practices.

---

[5]Total variable cost $m + l_i$ is pivotal. Alternatively, one can assume that firms differ in their variable cost due to exogenous reasons. I attribute the difference in the variable cost to the difference in transportation cost to match the empirical evidence.

[6]This subsection relies heavily on the note in Arrow (1969) and the empirical evidence in Teece (1977).

[7]Some studies that consider the adoption of CNC are Kelley and Helper (1999), Baptista (2000), and Stoneman (2002).

Finally, note that learning from interactions has a temporal dimension. Learning is only possible in case of sequential investment as the non-adopting firm learns from the adoption experience of their rival. Learning reduces adoption cost in the second period. Learning is impossible when both firms simultaneously invest. Table 1.1 summarizes the adoption costs in the two situations:

Situation	Investment Decision	Learning	**Adoption costs of**	
			firm i	**firm** j
If both firms adopt in the either period	Simultaneous	No	I_0	I_0
If firm i adopts in the first period and firm j adopts in the second period	Sequential	Yes	I_0	$(1-\kappa) I_0$

Table 1.1: Effect of learning on adoption cost

$\kappa \in (0,1)$ represents learning benefits. Learning lowers the second period adoption cost by a factor $1 - \kappa$. For example, if the firms are geographically close to each other, then $\kappa >> 0$, which implies that firm j—which may adopt after firm i—learns from the adoption experience of firm i.

Timing and state variables that determine Cournot quantities in each period

Figure 1.1 shows the timing of the game in both periods. Each period involves two stages: the first stage concerns the adoption decision, and the second stage concerns production. Each firm observes the demand level X_t at the beginning of each period. Subsequently, each firm decides whether or not to adopt the new technology. Prior to production, each firm also observes its rival's adoption decision. Afterward, both firms produce quantities q_{it}. Formally, there are three state variables: demand level X_t, a discrete variable $\theta_i \in \{0, 1\}$, which takes a value of 1 if firm i adopts (and 0 otherwise) and another discrete variable $\theta_j \in \{0, 1\}$, which takes a value of 1 if firm j adopts (and 0 otherwise).

Mathematically, there is mapping from state variables $(X_t, \theta_i, \theta_j)$ to firm i's quantity q_i and earnings π_i:
$$(X_t, \theta_i, \theta_j) \mapsto q_i(X_t, \theta_i, \theta_j) \text{ and } \pi_i(X_t, \theta_i, \theta_j).$$

The following assumption regarding $\underline{X}$ ensures positive supply by each firm:

ASSUMPTION 1. $\underline{X} > 2l_2 - l_1 + m > 0.$
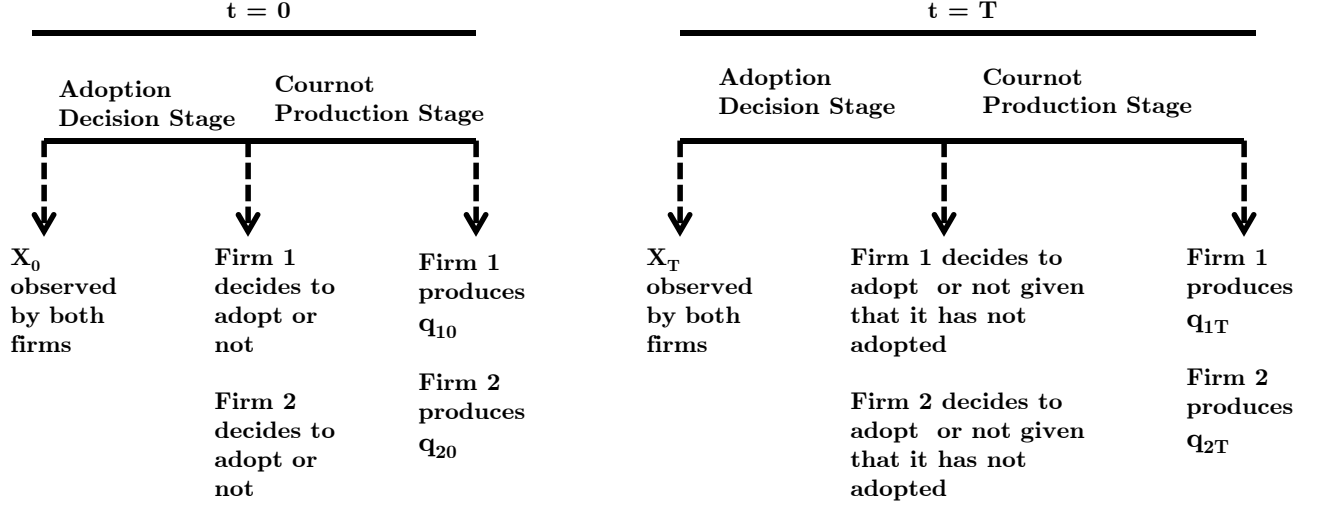
Figure 1.1: Timing of the game

Effect of Competition

Quantities and earnings per period

Assumption of linear inverse demand and constant marginal cost leads to closed form expressions for equilibrium quantities and earnings in each period. Cournot quantities are linear in demand level X_t:

$$q_i(X_t, \theta_i, \theta_j) = \frac{1}{3}\left[\underline{X} + VC_j(\theta_j) - 2VC_i(\theta_i)\right] + \frac{X_t}{3} \quad \text{for } i, j \in \{1, 2\}, \tag{1.1}$$

where

$$VC_i(0) = m + l_i \text{ and } VC_i(1) = m + l_i - \alpha \qquad \text{for } i \in \{1, 2\}$$

is the total variable cost of firm i. Earnings are a quadratic function of the demand level X_t:

$$\pi_i(X_t, \theta_i, \theta_j) = e_{i0}(\theta_i, \theta_j) + e_{i1}(\theta_i, \theta_j)X_t + e_{i2}(\theta_i, \theta_j)X_t^2, \tag{1.2}$$

where

$$e_{i0}(\theta_i, \theta_j) = \frac{e(\theta_i, \theta_j)^2}{9}; \; e_{i1}(\theta_i, \theta_j) = \frac{e(\theta_i, \theta_j)}{9}; \; e_{i2}(\theta_i, \theta_j) = \frac{1}{9}; \text{ and } e(\theta_i, \theta_j) \equiv \underline{X} + VC_j(\theta_j) - 2VC_i(\theta_i).$$

Upon inspection, quantity q_i and earnings π_i, depend on the relative transportation cost between firms $VC_i(\theta_i) - VC_j(\theta_j)$. The difference leads to a series of inequalities given in the following Lemma.

LEMMA 1. *Adoption by firm i negatively affects firm j:*

(i) **Quantity and earnings of firm i decreases as firm j adopts:**

$$k(X_t, 1, 0) > k(X_t, 1, 1) > k(X_t, 0, 0) > k(X_t, 0, 1) \, for \, k \in \{q_i, \pi_i\}.$$

(ii) **Supply increases when either firm adopts:**

$$q_i(X_t, 1, 1) + q_j(X_t, 1, 1) > q_i(X_t, 1, 0) + q_j(X_t, 0, 1) > q_i(X_t, 0, 0) + q_j(X_t, 0, 0).$$

(iii) **Increase in earnings is higher for firm i when it adopts first:**

$$\pi_i(X_t, 1, 0) - \pi_i(X_t, 0, 0) \; > \; \pi_i(X_t, 1, 1) - \pi_i(X_t, 0, 1).$$

Suppose firm 1 adopts in the first period. Firm 1's quantity produced increases from $q_1(X_0, 0, 0)$ to $q_1(X_0, 1, 0)$. In response, firm 2's quantity decreases from $q_2(X_0, 0, 0)$ to $q_2(X_0, 0, 1)$. In aggregate, the total supply increases from $q_1(X_0, 0, 0) + q_2(X_0, 0, 0)$ to $q_1(X_0, 1, 0) + q_2(X_0, 0, 1)$ and hence market clearing price decreases. Firm 1's earnings increase from $\pi_1(X_0, 0, 0)$ to $\pi_1(X_0, 1, 0)$ and firm 2's earnings decreases from $\pi_2(X_0, 0, 0)$ to $\pi_2(X_0, 0, 1)$. This means that some of the gain in earnings for firm 1 after adoption come at the expense of firm 2.

The feature that adoption has a negative effect on the rival is due to the assumption of Cournot competition. This result is consistent with Stoneman and Kwon (1996) who find that earnings of non-adopters decrease as the number of adopters increase. Generally, this result will arise as long as the goods produced by both firms are strategic substitutes.

Consider the effect of competition and location without learning. The last inequality implies that adoption gains depend on the order of adoption: Adopting first is better than adopting second. This leads both firms to adopt in the first period. Now, add learning. Consider the limiting case where the learning benefits are drastic so that $\kappa \approx 1$. Suppose firm i adopts in the first period. Firm j adopts in the second period for sure since adoption cost is zero due to drastic learning. Consequently, firm i does not remain the sole user of the new technology, which in turn reduces its incentive to adopt in the first period. Herein lies the fundamental tradeoff: competition induces firms to adopt early while learning induces firms to wait.

In the next section, I formally analyze the tradeoff by deriving equilibrium.

1.4 Equilibrium derivation

Equilibrium Overview

There are two interesting possibilities concerning the relative timing of investment in new technology. First, the relative timing can be sequential in nature in which firm i invests in

the first period and becomes the *leader*. Firm j, which is naturally the *follower*, has the option to invest in the second period. Second, the relative timing can be simultaneous in nature in which both firms simultaneously invest in either period.

Of course, there is a third possibility: Neither firm invests in the new technology in either period. Intuitively, both firms will never invest in new technology if the demand level in the first period is sufficiently low: $X_0 \le \underline{X_0}$. The expression for $\underline{X_0}$ is provided later in this section. Heuristically, the threshold $\underline{X_0}$ is the demand level at which it is just feasible for firm 1 to invest in the first period. This no-investment possibility is not interesting and so from here onwards in this deterministic setup, I assume that $X_0 > \underline{X_0}$.

The strategy of firm i involves adoption time denoted by τ_i and quantity q_i produced in both periods. Each firm must contemplate its value if it is the leader or the follower, or if it simultaneously invests in either period. Equilibrium strategies depend on the present value of earnings in each possibility. The equilibrium concept is that of sub-game perfect equilibrium, which rules out empty threats. In case of multiple equilibria that can be Pareto ranked, I assume that firms coordinate on the Pareto-Superior equilibrium.

Formally, a sub-game perfect equilibrium is a pair of strategies $[\tau_i^*, \{q_i^*\}_{t \in \{0,T\}}]$ for $i \in \{1,2\}$ such that each firm maximizes the present value of earnings for every demand level X_t given the equilibrium strategy of the rival.

Firm value as a follower

Suppose firm j has already invested in the new technology in the first period so that it becomes the leader. Firm i as a follower has the option to adopt in the second period.

Let $V_i^F(X_0)$ denote the value of the follower firm i. The follower's value consists of two components: The first component represents the present value of earnings from not adopting the new technology. This is the value from assets in place. The second component represents the growth option that denotes the increase in earnings when firm i adopts the new technology. Mathematically,

$$V_i^F(X_0) = W_F^i(X_0) + O_i^F(X_0), \tag{1.3}$$

where

$$W_F^i(X_0) = \underbrace{\pi_i(X_0, 0, 1)}_{\text{Firm } i\text{'s first period earnings}} + \underbrace{\pi_i(X_T, 0, 1)}_{\text{Firm } i\text{'s second period earnings}} e^{-rT} \tag{1.4}$$

is the value from assets in place and

$$O_F^i(X_0) \;=\; \max\left(\underbrace{\pi_i(X_T, 1, 1) - \pi_i(X_T, 0, 1)}_{\text{Firm } i\text{'s change in earnings}} - \underbrace{(1 - \kappa)I_0}_{\text{Adoption costs with learning}} ,\; 0 \right) e^{-rT}$$

$$\;=\; A_i^F\, e^{-rT}\, \max\left(X_T - X_i^F, 0 \right); \text{ and } X_i^F \equiv \frac{B_i^F + (1 - \kappa)I_0}{A_i^F} > 0. \qquad (1.5)$$

is the growth option that represents the increase in earnings when firm i adopts with

$$A_i^F \;=\; \frac{4\alpha}{9} > 0; \text{ and } B_i^F \;=\; \frac{4}{9}\alpha(2l_i - l_j + m - \underline{X}) < 0.$$

Investment thresholds X_i^F of both firms are:

$$X_2^F \;=\; 2l_2 - l_1 + m - \underline{X} + \frac{9I_0(1 - \kappa)}{4\alpha}; \text{ and } X_1^F \;=\; X_2^F - 3(l_2 - l_1). \qquad (1.6)$$

Upon inspection of equation (1.5), follower firm i is long A_i^F number of calls with a strike price of X_i^F. Also note that the follower threshold for firm 1 (X_1^F) is lower than the follower threshold for firm 2 (X_2^F).

The following technical assumption regarding the adoption cost in the second period ensures that the investment threshold for firm 1 as a follower is positive.

ASSUMPTION 2. *Adoption cost I_0 is sufficiently high: $I_0 \geq \frac{4\alpha(l_2 - 2l_1 - m + \underline{X})}{9(1 - \kappa)}$.*

Comparative statics of firm 2's investment threshold is given in the following lemma.[8]

LEMMA 2. *The investment threshold of follower firm 2, X_2^F*

  (i) *increases if relative transportation cost $(l_2 - l_1)$ increases,*

  (ii) *increases if efficiency of the new technology (α) increases,*

  (iii) *decreases if learning benefits (κ) increases.*

As the relative transportation cost $l_2 - l_1$ increases, the quantity produced by firm 2 decreases. Therefore, it has less incentive to adopt a variable cost reducing technology since the adoption

---

[8]I focus on comparative statics of follower firm 2's investment threshold as firm 1 will never be the follower in equilibrium. The qualitative features of the comparative statics also hold for X_1^F.

costs are sunk ($(1-\kappa)I_0 >> 0$). Similarly, an increase in the efficiency of new technology α increases quantity produced by firm 1. In response, quantity produced by firm 2 decreases, which again lowers the incentive of firm 2 to adopt in the second period. Finally, the adoption cost decreases for firm 2 since it learns from the adoption experience of firm 1.

The level of firm 2's investment threshold is further understood by comparing it with the monopoly threshold as shown below.

Effect of learning and market structure

Suppose firm 2 is a monopolist. Standard calculations yield that per period earnings with and without adoption are

$$\pi_2^M(X_t,1) = \tfrac{1}{4}\left(X_t + \underline{X} - m - l_2 + \alpha\right)^2 \quad \text{and} \quad \pi_2^M(X_t,0) = \tfrac{1}{4}\left(X_t + \underline{X} - m - l_2\right)^2$$

respectively. Similar to the analysis above, firm 2 adopts if earnings from adoption exceed earnings from non-adoption. Intuitively, firm 2 adopts if the demand level X_T in the second period is sufficiently high. This condition is equivalent to $X_T > X_2^M$ where

$$X_2^M \equiv \frac{-\alpha}{2} + l_2 + m + \frac{2I_0}{\alpha} - \underline{X}. \tag{1.7}$$

Comparing equations (1.6) and (1.7) yields

$$X_2^F - X_2^M = \frac{\alpha}{2} + l_2 - l_1 + \frac{I_0(9(1-\kappa)-8)}{4\alpha}. \tag{1.8}$$

Figure 1.2 plots $X_2^F - X_2^M$ as a function of the learning benefits κ for low and high values of relative transportation cost $l_2 - l_1$. Increases in learning decreases $X_2^F - X_2^M$ while increases in relative transportation cost increases $X_2^F - X_2^M$.

When firms are geographically far from each other so that (i) learning benefits are low ($\kappa \approx 0$) and (ii) relative transportation costs are high (high $l_2 - l_1$), firm 1 effectively raises the adoption cost of firm 2 by investing in the first period. This result is similar in spirit to Salop and Scheffman (1983), who indicate the use of predatory pricing in raising rivals' costs. Equation (1.8) points to an alternate mechanism: First period investment by firm 1 raises second period adoption cost for firm 2. On the other hand, when firms are geographically close, learning reduces the adoption cost of firm 2.

Proposition 1 summarizes the economic reasoning behind equation (1.8).
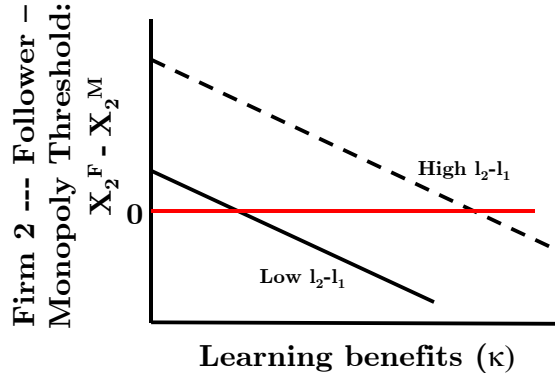
Figure 1.2: This figure shows the relationship between difference in firm 2's investment threshold as a follower and monopoly $(X_2^F - X_2^M)$ with learning benefits (κ) and transportation cost $l_2 - l_1$. The relationship is given in equation (1.8).

PROPOSITION 1. *If firms are geographically far from each other, then the investment threshold of follower firm 2 in a duopoly is higher than the investment threshold of firm 2 in a monopoly. This result reverses if firms are geographically close.*

Firm 1 as a leader needs to evaluate the benefits of adopting in the first period taking the negative impact of learning into account. The next section, which precisely calculates incentives of both firms to become the leader, formalizes this notion.

Firm value as a leader

Let $V_i^L(X_0)$ denote the value of the leader firm i after it adopts in the first period, conditional upon follower firm j pursuing its optimal exercise strategy. The leader's value consists of two components: The first component represents the present value of earnings assuming follower firm j never adopts. This is the value from assets in place. The second component represents the loss in earnings in the second period when firm j adopts. Mathematically,

$$V_i^L(X_0) = W_i^L(X_0) + O_i^L(X_0), \tag{1.9}$$

where

$$W_i^L(X_0) = \underbrace{\pi_i(X_0, 1, 0)}_{\text{Firm } i\text{'s first period earnings}} + \underbrace{\pi_i(X_T, 1, 0)}_{\text{Firm } i\text{'s second period earnings}} e^{-rT} - I_0 \tag{1.10}$$

is the value from assets in place and

$$O_i^L(X_0) \;=\; \left[\underbrace{\mathbf{1}(X_T \geq X_j^F)}_{\text{Event when firm } j \text{ adopts}} \left(\underbrace{\pi_i(X_T, 1, 1) - \pi_i(X_T, 1, 0)}_{\text{Loss in earnings when firm } j \text{ adopts}} \right) \right] e^{-rT},$$

$$= \; A_i^L \left[\mathbf{1}(X_T \geq X_j^F)\left(X_T - X_i^L\right) \right] e^{-rT}; \quad X_i^L \equiv \frac{B_i^L}{A_i^L} \tag{1.11}$$

represents loss in earnings if firm j adopts with

$$A_i^L \;=\; \frac{-2\alpha}{9} \;<\; 0; \text{and } B_i^L \;=\; \frac{1}{9}\alpha(3\alpha - 4l_j + 2l_i - 2m + 2\underline{X}) \;>\; 0.$$

Firm i as a leader wants the demand level in the first period X_0 to increase[9] but it does not want the demand to increase too much. Firm j adopts in the second period if the demand increases too much, which reduces the earnings of the leader firm i. This is akin to firm i being short a call option. Specifically, firm i is short firm j's growth option. However, the option is not a typical call option since the exercise event $\{X_T > X_j^F\}$ is different from the exercise price X_i^L.

Increase in learning benefits reduces the investment threshold of follower firm j, which reduces O_i^L. The following lemma summarizes the effect of learning on the value of being the leader.

LEMMA 3. *Learning decreases the present value of earnings from being the leader.*

The next section describes how competition may lead to early investment.

Effect of competition on early adoption

There are two different cases. First, suppose the first period demand level $X_0 = x$ where x is such that firm 2 prefers being the leader over being the follower, i.e.,

$$V_2^L(x) - V_2^F(x) \;>\; 0.$$

Since firm 1 has a lower transportation cost than firm 2, it has to be the true that firm 1 also prefers being the leader over being the follower, i.e.,

$$V_1^L(x) - V_1^F(x) \;>\; 0.$$

That is both firms find it beneficial to be leader. Therefore, firm 1 decides to invest at a slightly lower threshold $x - \epsilon$. In turn, firm 2 decides to invest at $x - 2\epsilon$. This process

---

[9]Recall that the demand level in the second period $X_T = X_0 e^{\mu T}$.

continues up to a threshold X_p^{21}, which is the point at which firm 2 is indifferent between being the leader or the follower. Formally, X_p^{21} is the solution to

$$V_2^L(y) - V_2^F(y) = 0. \tag{1.12}$$

In the appendix, I show that $X_p^{21} \in (0, X_2^F)$ and it is unique.

Note that at X_p^{21}, firm 1 is still better off being the leader or being the follower, i.e.,

$$V_1^L(X_p^{21}) - V_1^F(X_p^{21}) > 0.$$

Now consider the second case where both firms acts myopically. They ignore their rival's response. Each firm optimally invests at $X_i^{LN} \equiv X_j^F e^{-\mu T} - \epsilon$. Each firm invests at a point where it knows that their rival as a follower does not invest in the second period. Since $X_1^F < X_2^F$, $X_2^{LN} < X_1^{LN}$.

Finally consider the best responses that combine these two cases. In the event $\{X_1^{LN} < X_p^{21}\}$, then firm 2 does not want to be the leader anyway, and firm 1 invests at $X_0 = X_1^{LN}$ without any fear of being preempted by firm 2. In the complementary event $\{X_1^{LN} \geq X_p^{21}\}$, firm 2 wants to the leader. But from the argument above, due to competition, the investment threshold decreases to X_p^{21}.

To summarize, firm 2 will never be the leader. Firm 1 invests in the first period at a threshold

$$X_0 \geq \min\{X_1^{LN}, X_p^{21}\}.$$

Lastly, each firm compares its sequential investment possibility with the simultaneous investment possibility, which is shown next.

Firm value with simultaneous investment

Simultaneous investment in the second period

Let $V_i^S(X_0)$ denote the value of firm i simultaneously investing with firm j in the second period. The simultaneous investment value consists of two components: The first represents the present value of earnings when neither firm invests in the new technology. This is the value from assets in place. The second component represents the growth option, which denotes the increase in earnings when both firms simultaneously invest in the new technology. Mathematically,

$$V_i^S(X_0) = W_i^S(X_0) + O_i^S(X_0), \tag{1.13}$$

where

$$W_i^S(X_0) = \underbrace{\pi_i(X_0,0,0)}_{\text{Firm } i\text{'s first period earnings}} + \underbrace{\pi_i(X_T,0,0)}_{\text{Firm } i\text{'s second period earnings}} e^{-rT} \qquad (1.14)$$

is the value from assets in place and

$$O_i^S(X_0) = \max\left(\underbrace{\pi_i(X_T,1,1) - \pi_i(X_T,0,0)}_{\text{Simultaneous investment change in earnings}} - \underbrace{I_T}_{\text{Adoption costs}}, 0 \right) e^{-rT}$$

$$= A_i^S e^{-rT} \max\left(X_T - X_i^S, 0\right); \quad X_i^S \equiv \frac{B_i^S + I_0}{A_i^S} > 0. \qquad (1.15)$$

is the growth option that represents the increase in earnings when firm i simultaneous adopts with

$$A_i^S = \frac{2\alpha}{9} > 0; \quad B_i^S = \frac{-1}{9}\alpha(\alpha - 4l_i + 2l_j - 2m + 2\underline{X}) < 0.$$

The investment threshold of firm i simultaneously investing with firm j, X_i^S, is

$$X_2^S = \frac{-\alpha}{2} - l_1 + 2l_2 + m + \frac{9I_0}{2\alpha} - \underline{X} \quad \text{and} \quad X_1^S = X_2^S - 3(l_2 - l_1). \qquad (1.16)$$

Upon inspection of equation (1.15), firm i is long A_i^S number of calls with a strike price of X_i^S.

Comparing the follower investment threshold for both firms in equation (1.6) with the simultaneous investment threshold for both firms in equation (1.16) yields a series of inequalities as given in the following lemma.

LEMMA 4. *The follower and simultaneous investment threshold are ranked as follows:*

$$0 < X_1^F < X_2^F < X_1^S < X_2^S$$

Due to the higher transportation cost of firm 2, i.e., $l_2 > l_1$, it is intuitive that $X_1^F < X_2^F$ and $X_1^S < X_2^S$. The reason why $X_2^F < X_1^S$ is subtle. Market clearing price decreases as any firm adopts (lemma 1). Thus, when both firms simultaneously adopt, there is a significant drop in price. In response, both firms simultaneously invest when the demand level in the second period is really high.

In the case of simultaneous investment, both firms have to coordinate investing at either firm 1's investment threshold X_1^S or firm 2's investment threshold X_2^S. The following lemma shows that firm 2 simultaneously invests with firm 1 at firm 1's investment threshold X_1^S.

LEMMA 5. *Firm* 2 *prefers simultaneously investing with firm* 1 *over being the follower, i.e.,*

$$V_2^S(X_0) > V_2^F(X_0) \qquad \forall X_0 \geq 0.$$

Proof. Three different cases arise:

Case 1: $X_T < X_2^F$ — In this case, the growth options are worthless: $O_2^F = O_2^S = 0$. From Lemma 1, it is clear that $W_2^S > W_2^F$.

Case 2: $X_2^F < X_T < X_1^S$ — In this case, follower firm 2's growth option is positive and the simultaneous growth option is worthless: $O_2^F > 0$ and $O_2^S = 0$. Follower firm 2's value V_2^F increases with the demand level X_0. Then the difference $V_2^S(X_0) - V_2^F(X_0)$ is at a minimum when learning benefits are drastic $\kappa \approx 1$ and when $X_0 = X_1^S e^{-\mu T}$:

$$V_2^S(X_1^S e^{-\mu T}) - V_2^F(X_1^S e^{-\mu T})\Big|_{\kappa=1} = -\frac{(e^{\mu T} - e^{rT})(4\alpha^2 + 12\alpha(l_2 - l_1)) + 4a\alpha(e^{\mu T} - e^{rT})}{\alpha} > 0,$$

where $\underline{X} = 2l_2 - l_1 + m + \alpha + a$ with $a > 0$ (Assumption 1).

Case 3: $X_T > X_1^S$ — In this case, both growth options are positive in value: $O_2^F = O_2^S > 0$. From Case 2 and $O_2^S > 0$, it has to be that $V_2^S(X_0) > V_2^F(X_0)$. □ $\qquad$ □

COROLLARY 1. *Both firms simultaneously invest in the second period in the event* $\{X_T > X_1^S\}$.

Proof. This is a direct implication of Lemmas 4 and 5. $\qquad$ □

If firm 1 decides to invest simultaneously, it will invest in the new technology in the second period if $X_T > X_1^S$. Firm 1 will invest with the understanding that firm 2 follows suit. This understanding is implicit—there is no contractual obligation that enforces firm 2 to simultaneously invest. Therefore firm 2 may deviate and invest in the first period. The implication of Lemma 5 is that firm 2 does not credibly deviate. If firm 2 deviates and invests in the first period, then firm 1 invests at a lower threshold in the first period, which makes firm 2 the follower and firm 2 does not want to be the follower.

Simultaneous investment in the first period

Intuitively, both firms adopt in the first period if the demand level is really high. This is equivalent to the condition that

$$\underbrace{\pi_2(X_0, 1, 1) + e^{-rT} \pi_2(X_T, 1, 1)}_{\text{Firm 2's value from simultaneous investment in the first period}} - I_0 > V_2^S(X_0).$$

Both firms simultaneously adopt in the first period in the event $\{X_0 > \overline{X_0}\}$ where $\overline{X_0}$ is the solution to

$$\pi_2(y, 1, 1) + e^{-rT} \pi_2(ye^{\mu T}, 1, 1) - I_0 - V_2^S(y) = 0.$$

Armed with investment thresholds and firm values in every single possibility, I derive equilibrium next.

Equilibrium selection

To summarize the optimal strategies, neither firm invests in the new technology if the demand level in the first period is sufficiently low, i.e., $X_0 < \underline{X_0}$. Also, both firms invest simultaneously in the first period if the demand level is sufficiently high, i.e., $X_0 > \overline{X_0}$. In the intermediate case, $X_0 \in (\underline{X_0}, \overline{X_0})$, two possibilities arise: First, firm 1 invests in the first period and firm 2 invests in the second period in the event $\{X_T > X_2^F\}$. Second, both firms simultaneously invest in the second period in the event $\{X_T > X_1^S\}$.

Figure 1.3 displays the normal form game with only dominated strategies when $X_0 \in (\underline{X_0}, \overline{X_0})$. The boxed values denote sub-optimal best responses. For example, $\boxed{V_2}$ in the top right panel denotes sub-optimal present value of earnings of firm 2. If firm 1 invests in the first period, the best response of firm 2 is to invest in the second period in the event $\{X_T > X_2^F\}$ (not $\{X_T > X_1^S\}$). Similarly, in the bottom left panel, the best response of firm 1 if firm 2 invests in the event $\{X_T > X_2^F\}$ is to invest in the first period. Sub-optimal firm value $\boxed{V_1}$ represents firm 1's present value of earnings if it invests in the second period in the event $\{X_T > X_1^S\}$.

I distinguish two cases:

- **Case A:** $V_1^L(X_0) > V_1^S(X_0)$;
- **Case B:** $V_1^L(X_0) \leq V_1^S(X_0)$.

First, consider Case A. The top left panel featuring sequential investment is the unique equilibrium. In this case, firm 1 invests first in the first period and firm 2 retains the option to invest in the second period.

Firm 2

Figure 1.3: This figure shows the firm value with only dominated strategies. The boxed values represent firm values with sub-optimal best responses.

Next, consider Case B. There are two pure strategy equilibria: Sequential investment in the top left panel and simultaneous investment in the bottom right panel. Any convex combination of these two equilibria is also an equilibrium — these combinations form the set of mixed strategy equilibria.

From Lemma 5, $V_2^S > V_2^F$ and since $V_1^S > V_1^L$, both firms are better off simultaneously investing with each other. Formally, SIM equilibrium Pareto dominates SEQ and all other mixed strategy equilibria. The next lemma summarizes this result.

LEMMA 6. *Simultaneous investment equilibrium, if it exists, Pareto dominates (from the firm's point of view) all other equilibria.*

In order to reduce the set of equilibria in Case B, I make the following assumption:

ASSUMPTION 3. (Equilibrium Selection) *Both firms simultaneously invest if* $V_1^L(X_0) < V_1^S(X_0)$.

Next proposition summarizes the investment strategies.

PROPOSITION 2. *The investment strategies of both firms are as follows:*

(i) *If* $X_0 \le \underline{X_0}$, *then neither firm invests in the new technology in either period,*

(ii) *If* $X_0 \ge \overline{X_0}$, *then both firms invests in the new technology in the first period,*

(iii) *If $X_0 \in (\underline{X_0}, \overline{X_0})$ and $V_1^L(X_0) > V_1^S(X_0)$, then firm 1 invests in the first period and firm 2 invests in the second period in the event $\{X_T > X_2^F\}$,*

(iv) *If $X_0 \in (\underline{X_0}, \overline{X_0})$ and $V_1^L(X_0) \leq V_1^S(X_0)$, then both firms invest in the second period in the event $\{X_T > X_1^S\}$.*

Comparative statics are given in the following lemma.

LEMMA 7. *The incentive of firm 1 to invest in the first period increases if*

(i) *learning benefits (κ) decrease:* $\dfrac{\partial(V_1^L - V_1^S)}{\partial \kappa} < 0$,

(ii) *relative transportation cost ($l_2 - l_1$) increases:* $\dfrac{\partial(V_1^L - V_1^S)}{\partial(l_2 - l_1)} > 0$,

(iii) *efficiency of new technology (α) increases:* $\dfrac{\partial(V_1^L - V_1^S)}{\partial \alpha} > 0$.

When firms are near each other, relative transportation cost $l_2 - l_1$ is low and learning benefits κ are high. Thus, firm 1's incentive to be the leader is low. On the other hand, learning does not impact firm value from simultaneous investment. Therefore, geographically close firms find it mutually beneficial to simultaneously invest. Observationally, technology adoption is geographically clustered and lag in adoption timing between the two firms is low.

The following two corollaries elaborate the combined effect of learning and competition.

COROLLARY 2. *With learning, competition may not cause firms to invest early.*

A well-known notion in the real options literature is that competition erodes the time value of waiting.[10] Corollary 2 places caveats on this notion. With simultaneous investment, both firms actually invest at a higher threshold X_1^S. Observationally, neither firm adopts for a long time, but when they adopt, they both adopt together.

COROLLARY 3. *If learning benefits are high (when firms are close to each other), then firms simultaneously invest which features no learning*

---

[10] Grenadier (2002) and Back and Paulsen (2009) show that the investment threshold approaches the positive NPV threshold as the intensity of competition increases. They assume singular investment costs but do not place restrictions on capacity increases from the new option. On the other hand, I place a restriction on the capacity increase due to new technology (only one growth option) but relax the assumption of singular investment costs. The fact that competition does not erode the option value to wait in a competitive setting has also been shown in Fudenberg and Tirole (1985), Pawlina and Kort (2006), Novy-Marx (2007), Carlson et al. (2011), and Bustamante (2011).
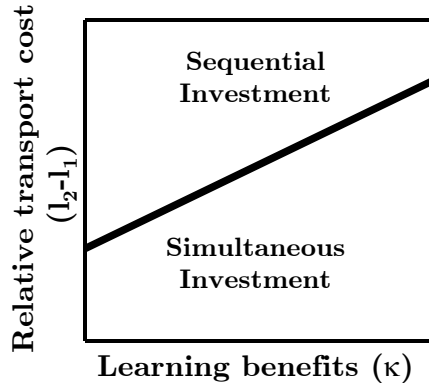
Figure 1.4: This figure shows equilibrium as a function of relative transport cost $l_2 - l_1$ and learning benefits κ.

Learning serves a dual role. In sequential investment, firm 2 learns from the adoption experience of firm 1. This is the learning in the traditional sense. However, learning reduces the value of the leader. The threat of learning induces firm 1 to simultaneously invest and as a result, neither firm learns. Figure 1.4 shows this counterintuitive result graphically.

1.5   Empirical Implications

Now, I summarize the empirical implications of the model, relating them to existing evidence.

(i) **Positive effect of learning from interactions on technology adoption:** The magnitude of learning from interactions leads to two different equilibria: First high learning benefits lead to simultaneous investment— the lag in adoption timing between firms is zero. Second, low learning benefits lead to sequential investment—learning decreases the follower's threshold. Therefore, unequivocally, learning reduces lag in adoption timing. Young (2009) provides positive evidence of social learning in technology adoption. Caselli and Coleman (2001), Keller (2002), Comin and Hobijn (2004), and Comin and Hobijn (2010) find that trade and foreign direct investment (activities which increase contact with foreign persons) increase the probability of technology adoption.

(ii) **Positive effect of size on technology adoption:** The model predicts a positive correlation between size and probability of technology adoption, which is corroborated in Karshenas and Stoneman (1993), Stoneman and Kwon (1996), and Baptista (2000). This idea goes back at least to Schumpeter [see the account in Hall and Khan (2003)]. He argued that larger firms adopt early since the benefits are greater for them. In the model, the variable cost of firm 1 is lower than that of firm 2. Therefore, it produces

more and hence the benefits of adopting a variable cost reducing technology is higher for firm 1.

(iii) **Geographic clustering of technology adoption:** The model predicts simultaneous investment when firms are geographically close to each other. This is corroborated in Kelley and Helper (1999), Baptista (2000), and No (2008).

(iv) **Effect of location on technology adoption:** In an international study of technology adoption across various industries, Keller (2002), Comin and Hobijn (2004), Comin and Hobijn (2010), Comin, Dmitriev, and Rossi-Hansberg (2012) find significant regional asymmetries. Firms located in countries that are closer to "adoption leaders" invest earlier in the new technology. Technology slowly trickles down to firms less well situated.

(v) **Effect of adopting first:** The increase in earnings from adopting first is higher than the increase in earnings from adopting second. This is corroborated in Karshenas and Stoneman (1993), Stoneman and Kwon (1996), and Baptista (2000).

These implications provide direct evidence of technology adoption. Since, location and learning affect investment strategies in the new technology, they also affect asset prices. This is next.

1.6   A stochastic model of technology adoption

Since investment in the new technology is a growth option held by both firms, location and learning also affect asset prices. In this section, I highlight how location and learning affect the risk profiles of firms.

Environment

Demand level shocks follow geometric Brownian motion[11]:

$$\frac{dX_t}{X_t} = \mu dt + \sigma d\mathbb{W}_t. \tag{1.17}$$

Furthermore, there is a unique stochastic discount factor whose dynamics are

$$\frac{dM_t}{M_t} = -rdt - \lambda d\mathbb{W}_t, \tag{1.18}$$

where $\lambda \equiv \frac{g-r}{\sigma}$ is the market price of risk.

---

[11]Uncertainty is modeled in the following way. The probability space is denoted by $(\Omega, \mathcal{F}, \mathbb{P})$ where the filtration $(\mathcal{F}_t)_{t \geq 0}$ represents the information resolved over time and is known to both the firms.

Game set up

The dynamic game is a repeated version of the game in Figure 1.1. After observing X_t at time t, both of firms decide whether to adopt or not adopt. This decision is public knowledge—both firms know their rival's adoption status also. Afterward, firms receive Cournot earnings, which is followed by realization of X_{t+dt} and the game is repeated.

I formalize the game below. Denote the first time when firm i adopts by τ_i and define

$$D_i(t) = \begin{cases} 1 \text{ if } t \geq \tau_i \\ 0 \text{ if } t < \tau_i. \end{cases}$$

The function $D_i(t)$ is the Heaviside step function which takes a value of 1 when firm i adopts. At each instant, both firms make their investment decision knowing the complete history of the game $\Phi_t = ([X_s]_{s \leq t}, [D_1(s), D_2(s)]_{s \leq t})$, which is common to both the players. Given the Markov nature of the environment, it is natural to restrict attention to the most recent state variables X_t[12] and $\boldsymbol{D(t^-)} \equiv [D_1(t^-), D_2(t^-)]$ where $D_i(t^-) \equiv \lim_{s \uparrow t} D_i(s)$.

The firm i value is

$$
\begin{aligned}
M_t V_i(X_t) &= \mathbb{E}_t \left[\int_t^\infty M_s \times \pi_i(X_s, \theta_i, \theta_j) \times \mathbf{1}[D_i(s) = \theta_i, D_j(s) = \theta_j] \, ds \right] \\
&\quad - \mathbb{E}_t \left[M_{\tau_i} \times I_0 \times ((1 - \kappa)\mathbf{1}[D_i(\tau_i) = 0, D_j(\tau_i) = 1]) \right] \\
&\quad + \mathbb{E}_t \left[\mathbf{1}[D_i(\tau_i) = 1, D_j(\tau_i) = 1] \right].
\end{aligned}
\tag{1.19}
$$

The first term in the integral is the present value of the flow of earnings and the second term is the present value cost under either sequential or simultaneous investment. I further assume that

$$\delta > \frac{r + \sigma^2}{2};$$

where $\delta \equiv g - \mu$ to ensure finite firm value.

A pure strategy Markov-perfect equilibrium is a pair of strategies $(\tau_i^*, \{q_i^*\}_{s \geq t})$, for $i \in \{1, 2\}$ such that each firm maximizes their firm value in equation (2.10) for every state $[X_t, \boldsymbol{D(t^-)}]$ given the equilibrium strategy of the rival.

Similar to the previous section, I use backward induction to derive the equilibrium of the game.

---

[12]Since X_t is known, M_t is known also.

Firm value as a follower

Assuming firm j has already adopted, present value of firm i at time t is

$$
\begin{aligned}
M_t V_i^F(X_t) &= \sup_{\tau_i^F \in \mathrm{T}} \mathbb{E}_t \left[\int_t^{\tau_F^i} \pi_i(X_s, 0, 1) \, M_s \, ds + \int_{\tau_i^F}^{\infty} \pi_i(X_s, 1, 1) \, M_s \, ds - M_{\tau_i^F}(1 - \kappa) I_0 \right] \\
&= W_i^F + O_i^F.
\end{aligned}
\tag{1.20}
$$

where $\tau_F^i = \min\{t : X_t \geq X_F^i\}$ and T is the set of stopping times. In the appendix, I show that

$$
W_i^F(X_t) = \frac{e_{i0}(0, 1)}{r} + \frac{X_t \, e_{i1}(0, 1)}{\delta} + \frac{X_t^2 \, e_{i2}(0, 1)}{2\delta - r - \sigma^2}
$$

is the value from assets in place and

$$
O_i^F(X_t) = \begin{cases}
\left(\dfrac{X_t}{X_i^F} \right)^{\gamma} \left(\dfrac{A_i^F X_i^F}{\delta} - \dfrac{B_i^F}{r} - (1 - \kappa) I_0 \right) & \text{if } X_t \leq X_i^F \\[2ex]
\dfrac{A_i^F X_t}{\delta} - \dfrac{B_i^F}{r} - (1 - \kappa) I_0, & \text{if } X_t > X_i^F
\end{cases}
$$

is the value of the growth option where

$$
X_i^F = \frac{\frac{B_i^F}{r} + (1 - \kappa) I_0}{\frac{A_i^F}{\delta}} \frac{\gamma}{\gamma - 1}; \quad \gamma = \frac{1}{\sigma^2} \left[-(r - \delta - \frac{1}{2}\sigma^2) + \sqrt{(r - \delta - \frac{1}{2}\sigma^2)^2 + 2r\sigma^2} \right] > 1.
$$

Firm value as a leader

Present value of firm i upon immediate investment at time $t = \tau_i^L$ is

$$
\begin{aligned}
M_t V_i^L(X_t) &= \mathbb{E}_t \left[\int_t^{\tau_F^j} \pi_i(X_s, 1, 0) \, M_s \, ds + \int_{\tau_j^F}^{\infty} \pi_i(X_s, 1, 1) \, M_s \, ds \right] - I_0 \\
&= W_i^L + O_i^L.
\end{aligned}
\tag{1.21}
$$

In the appendix, I show that

$$
W_i^L(X_t) = \frac{e_{i0}(1, 0)}{r} + \frac{X_t \, e_{i1}(1, 0)}{\delta} + \frac{X_t^2 \, e_{i2}(1, 0)}{2\delta - r - \sigma^2} - I_0
$$

is the value from assets in place and

$$
O_i^L(X_t) = \begin{cases}
\left(\dfrac{X_t}{X_j^F} \right)^{\gamma} \left(\dfrac{A_i^L X_j^F}{\delta} - \dfrac{B_i^F}{r} \right) & \text{if } X_t \leq X_j^F \\[2ex]
\dfrac{A_i^L X_t}{\delta} - \dfrac{B_i^L}{r}, & \text{if } X_t > X_j^F
\end{cases}
$$

is the loss in earnings when firm j adopts.

It is also useful to consider optimal investment strategy of firm i when it knows that it will be the leader. This is the hypothetical value when firm j credibly pre-commits to be the follower. Formally,

$$
\begin{aligned}
M_t \, V_i^{LN}(X_t) \;\; = \;\; & \sup_{\tau_i^{LN} \in \mathrm{T}} \mathbb{E}_t \left[\int_t^{\tau_i^{LN}} \pi_i(X_s, 0, 0) \, M_s ds \;+\; \int_{\tau_i^{LN}}^{\tau_j^F} \pi_i(X_s, 1, 0) \, M_s \, ds \right] \\
& + \;\; \mathbb{E}_t \left[\int_{\tau_j^F}^{\infty} \pi_i(X_s, 1, 1) \, M_s \, ds \right] \\
& - \;\; \mathbb{E}_t \left[M_{\tau_i^{LN}} I_0 \right] \\
& = \;\; W_i^{LN} + O_i^{LN}
\end{aligned}
\tag{1.22}
$$

where $\tau_i^{LN} = \min\{t : X_t \geq X_i^{LN}\}$. In the appendix, I show that

$$
W_i^{LN}(X_t) = \frac{e_{i0}(0,0)}{r} + \frac{X_t \, e_{i1}(0,0)}{\delta} + \frac{X_t^2 \, e_{i2}(0,0)}{2\delta - r - \sigma^2}
$$

is the value from assets in place and

$$
O_i^{LN}(X_t) = \begin{cases} \left(\dfrac{X_t}{X_i^{LN}} \right)^{\gamma} V_i^L(X_t), & \text{if } X_t \leq X_i^{LN} \\[3mm] V_i^L(X_t), & \text{if } X_t > X_i^{LN} \end{cases}
$$

is the value of the growth option where

$$
X_i^{LN} = \frac{B_i^{LN} + I_0}{A_i^{LN}} \frac{\theta}{\theta - 1}; \quad A_i^{LN} = A_i^S - A_i^L; \quad B_i^{LN} = B_i^S - B_i^L.
$$

Firm value with simultaneous investment

The present value of firm i when it simultaneously invests with firm j is

$$
\begin{aligned}
M_t \, V_i^S(X_t) \;\; = \;\; & \sup_{\tau_i^S \in \mathrm{T}} \mathbb{E}_t \left[\int_t^{\tau_i^S} \pi_i(X_s, 0, 0) \, M_s \, ds + \int_{\tau_i^S}^{\infty} \pi_i(X_s, 1, 1) \, M_s \, ds - M_{\tau_i^S} I_0 \right] \\
& = \;\; W_i^S + O_i^S.
\end{aligned}
\tag{1.23}
$$

where $\tau_i^S = \min\{t : X_t \geq X_i^S\}$. In the appendix, I show that

$$
W_i^S(X_t) = \frac{e_{i0}(0,0)}{r} + \frac{X_t \, e_{i1}(0,0)}{\delta} + \frac{X_t^2 \, e_{i2}(0,0)}{2\delta - r - \sigma^2}
$$

is the value from assets in place and

$$O_i^S(X_t) = \begin{cases} \left(\dfrac{X_t}{X_i^S}\right)^\gamma \left(\dfrac{A_i^S X_i^F}{\delta} - \dfrac{B_i^S}{r} - I_0\right), & \text{if } X_t \leq X_i^S \\[3ex] \dfrac{A_i^S X_t}{\delta} - \dfrac{B_i^S}{r} - I_0, & \text{if } X_t > X_i^S \end{cases}$$

is the value of the growth option from simultaneous investment where

$$X_i^S = \frac{\frac{B_i^S}{r} + I_0}{\frac{A_i^S}{\delta}} \frac{\gamma}{\gamma - 1}.$$

Equilibrium properties

For simplicity, assume that the demand level X_t is low so that neither firm wants to invest immediately. The qualitative properties of the equilibrium remain the same as in Proposition 2.

First consider the properties in sequential investment. Without any fear of preemption, firm 1 optimally invests at τ_i^{LN}. On the other hand, if $V_2^L(X_i^{LN}) - V_2^F(X_i^{LN}) > 0$, then firm 2 finds it beneficial to be the leader over being the follower. Therefore, firm 2 preempts and the investment threshold decreases to X_P^{21} (equation (1.12)) at which point, firm 2 is indifferent between being the leader or the follower. Therefore, firm 1 invests at the first time demand level reaches $\min\{X_i^{LN}, X_p^{21}\}$.[13]

Second consider the properties in simultaneous investment. Since transportation cost of firm 1 is lower than that of firm 2, the only candidate for a simultaneous investment threshold is X_1^S. For simultaneous investment to occur, Firm 1's value as a leader has to be lower than firm 1's value from simultaneous investment, i.e.,

$$V_1^S(x) > V_1^L(x) \quad \forall x \in (0, X_1^S).$$

Otherwise, firm 1 invests when the demand level reaches $\min\{X_i^{LN}, X_p^{21}\}$.

In the appendix, I show that when firms are geographically close, then both firms simultaneously invest. Observationally, there is not a significant lag between adoption timing of both firms. On the other hand, when firms are geographically distant, then investment in the new technology is sequential in nature. Firm 1—which is in a better location—invests at the first time demand level reaches $\min\{X_i^{LN}, X_p^{21}\}$. Firm 2 invests at the first time demand level reaches X_2^F. Observationally, there is a significant lag between adoption timing of both firms. The next proposition summarizes the equilibrium.

---

[13]If $X_i^{LN} < X_p^{21}$, firm 1 invests without fear of any preemption by firm 2. Carlson et al. (2011) label this equilibrium as *Leader-Follower non-preemptive equilibrium*. On the other hand, if $X_i^{LN} > X_p^{21}$, firm 1 fears preemption by firm 2. Carlson et al. (2011) label this equilibrium as *Leader-Follower preemptive equilibrium*.
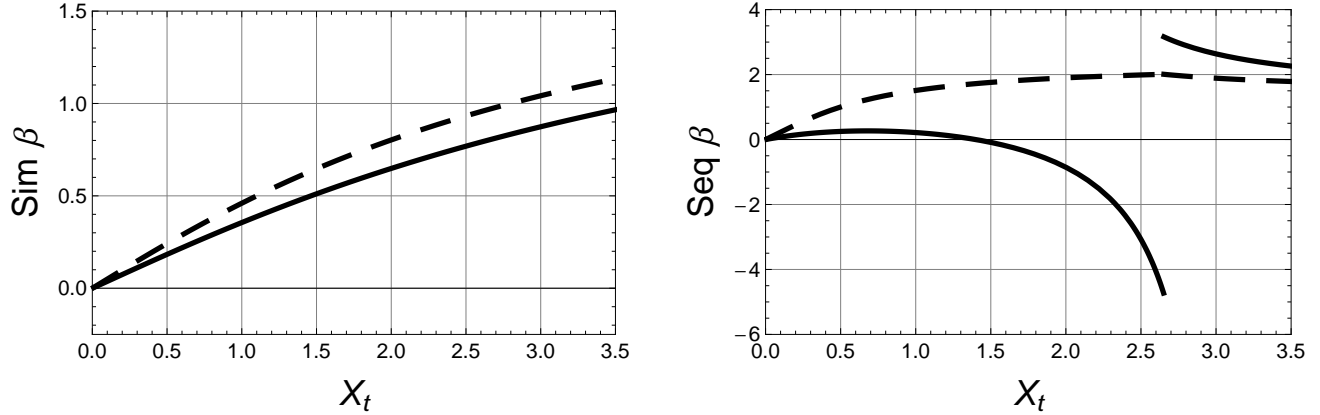
Figure 1.5: The left panel plots firms' βs from simultaneous investment. The right panel plots firms' βs from sequential investment assuming firm 1 has already invested. Solid line is for firm 1 and dashed line is for firm 2. The parameters are $\alpha = 3.5$, $m = 4$, $l_1 = 0.5$, $l_2 = 1$, $r = 0.1$, $\delta = 0.09$, $I_0 = 150$, $\kappa = 0.65$, and $\sigma = 0.2$

PROPOSITION 3. *The Markov perfect equilibrium is characterized by*

(i) ***Simultaneous equilibrium:*** *If firms are sufficiently close to each other, i.e. $l_2 - l_1 < l^*$, then both firms simultaneously adopt at τ_1^S.*

(ii) ***Sequential equilibrium:*** *If firms are sufficiently distant from each other, i.e. $l_2 - l_1 > l^{**}$, then firm 1 adopts at τ_1^{LN} and firm 2 adopts at τ_2^F. There is no credible preemption attempt by firm 2. In the intermediate case, i.e. $l^* \leq l_2 - l_1 \leq l^{**}$, firm 1 adopts at τ_p^{21} where $\tau_p^{21} = \min\{t : X_t \geq X_p^{21}\}$. Firm 1 adopts earlier than optimal due to the fear of preemption by firm 2. Firm 2 adopts at τ_2^F.*

In the next section, I analyze how location and learning affect firm risk.

Effect of geography on firms' risk loadings (βs)

Following Carlson, Fisher, and Giammarino (2004), the dynamic betas of each firm are

$$\beta_i(X_t) = \frac{\partial V_i(X_t)}{\partial X_t} \frac{X_t}{V_i(X_t)}. \tag{1.24}$$

Upon inspection, risk loading (β) of each firm depends on the germane equilibrium as given in Proposition 3, which in turn depends on the relative transportation cost $l_2 - l_1$. In this manner, geography affects asset pricing.

Left panel in Figure 1.5 plots the relationship between the firm β and the demand level X_t under simultaneous investment while the right panel plots the relationship under sequential investment. One result is clear: the risk dynamics are markedly different for different equilibria. Consider the risk loading of firm 1 in SEQ equilibrium after it has already adopted. Any positive news (higher values of X_t) increases the probability that follower firm 2 adopts. Therefore, positive news in the product market is dampened by the competitors growth option—competition acts like a natural hedge. The effect of competition is enough so that the risk loadings of both firms correlate negatively. In the SIM equilibrium, positive news in the product market affects both firms in the same direction. This leads to the following proposition.

PROPOSITION 4. *When firms are geographically close, their risk loadings correlate positively.*

Proof. This follows directly from Proposition 3. When firms are geographically close, the relative transportation cost $l_2 - l_1 < l^*$ and hence both firms simultaneously invest. $\quad\square$

Effect of geography on correlation between stock returns

Since conditional CAPM holds, firm i's expected return (from equation (2.10)) is

$$\mathbb{E}_t[R_i(X_t)] - rdt = \beta_i(X_t)\lambda; \quad \text{where } R_i(X_t) \equiv \frac{\pi_i(X_t, \theta_i, \theta_j)\,dt\,+\,dV}{V(X_t)}.$$

Assuming that realized returns also have an idiosyncratic component[14] with variance $\sigma_{\epsilon_i}^2$, correlation between the returns of both firms can be computed. Figure 1.6 plots correlation of returns for different equilibria. As expected, the correlations are markedly different for different equilibria. Under sequential investment (solid line), the correlation actually becomes negative while under simultaneous investment (solid line), correlation remains positive. This leads to the following proposition.

PROPOSITION 5. *When firms are geographically close, returns co-move together.*

Proof. This follows directly from Proposition 3. When firms are geographically close, the relative transportation cost $l_2 - l_1 < l^*$ and hence both firms simultaneously invest. $\quad\square$

---

[14]This is outside the model but incorporating the idiosyncratic piece is straightforward.
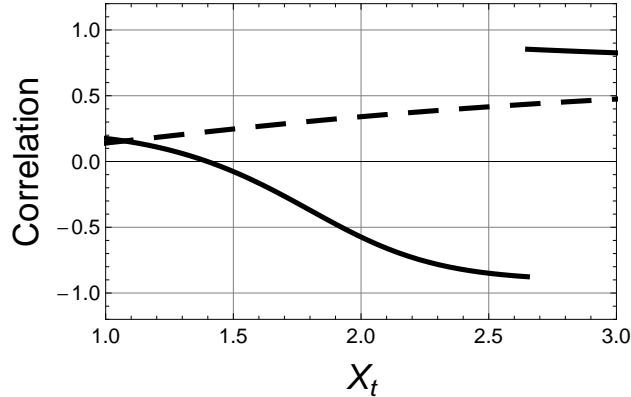
Figure 1.6: Solid line plots correlation under sequential investment and dashed line plots correlation under simultaneous investment. The parameters are $\alpha = 3.5$, $m = 4$, $l_1 = 0.5$, $l_2 = 1$, $r = 0.1$, $\delta = 0.09$, $I_0 = 150$, $\kappa = 0.65$, $\sigma_{\epsilon_1} = \sigma_{\epsilon_2} = 0.2$, and $\sigma = 0.2$.

Empirical implications of geography on asset pricing

Now, I summarize the empirical implications that concern asset pricing.

(i) **Local co-movement of stock returns:** When firms are geographically close, they simultaneously invest. Therefore, their βs' and correlation co-move together. Pirinsky and Wang (2006), Eckel et al. (2011), Barker and Loughran (2007), and Wongchoti and Wu (2008) provide suggestive evidence.

(ii) **Effect of competition on stock returns:** In concentrated industries, so that $l_2 - l_1$ is large, investment is sequential. The firms' βs do not co-move together. On the other hand, in less concentrated industries, so that $l_2 - l_1$ is small, both firms simultaneous invest. Then βs of firms co-move together. This is corroborated in Hoberg and Phillips (2010) and Bustamante (2011) who find that stock returns co-move together in less concentrated industries.

(iii) **Effect of geography on investment:** When firms are geographically close, they simultaneously invest. Dougal, Parsons, and Titman (2012) find that corporate investment is indeed geographic clustered.

1.7   Conclusion

This paper presents a model that explains the role of location and learning from interpersonal interactions on technology adoption. The setup involves two firms who have the option to adopt a new technology. While the logic of the positive role of learning is straightforward, the logic either explicitly or implicitly ignores imperfect competition. Imperfect competition

introduces strategic effects. One one hand, competition induces firms to adopt early and on the other hand, learning induces firms to wait. There are two equilibria that arise from the tradeoff— these equilibria depend on the relative location of the firms. First, firms located in better locations never adopt after firms less well situated. Observationally, there is geographic dispersion in technology adoption. Second, technology adoption is geographically clustered. Observationally, the lag in adoption timing is small for geographically close firms. To summarize, the spatial and temporal patterns of technology adoption are two sides of the same coin. Technology adoption diffuses through time and also through space.

Since location and learning affect the investment decision, they also affect asset prices. Specifically, the risk dynamics and returns of geographically close firms correlate positively. These predictions are testable, although one has to be careful since location choice by itself is endogenous. While the understanding of industrial organization on asset prices has been studied (Hoberg and Phillips (2010) and Bustamante (2011)), I am not aware of any studies that document the impact of geography on the cross section of stock returns and investment. Therefore, I anticipate future work testing the predictions of the model.

Chapter 2

Debt maturity and term spread

2.1 Introduction

Along with optimal leverage and seniority, firms also choose the time to maturity of their debt issuance. Overwhelming evidence (see Section 2.2) suggests that firms choose debt maturity based on the the term spread where term spread is defined as the difference between the 10-year Treasure note yield and the 90-day Treasury bill yield. Specifically, firms issue short term debt when the term spread is significantly positive and they increase maturity as the term spread decreases. That is, debt maturity and term spread are inversely related. In this paper, we provide a theoretical explanation of how term spread affects optimal debt maturity using the tradeoff theory of capital structure.

Our explanation is in contrast with the current theoretical literature that relies on the existence of either informational asymmetry, lack of liquidity, or agency conflicts. For instance, Flannery (1986) and Diamond (1991) show that good quality firms issue short-term debt as a signalling device to separate themselves from other poor quality firms when there is asymmetric information between debt investors and firm managers. Milbradt and He (2012) create a model framework where liquidity in the secondary market plays a central role in determining the optimal debt maturity of firms facing debt rollover risk. Myers (1977) and Johnson (2003) show that firms issue short-term debt to overcome the problem of under-investment created by agency conflicts. In our model, firms face the following tradeoff. On one hand, debt issuance provides tax benefit to the firms but on the other hand, debt issuance is also accompanied with higher bankruptcy and transaction costs.

We develop a dynamic capital structure model with stochastic interest rates to highlight that optimal debt maturity mainly depends on the tradeoffs between bankruptcy costs and transaction costs of debt rollover. A significantly positive term spread induces a higher risk of financial distress. This is because the firm's expected growth rate is lower than

the long run equilibrium growth rate. Therefore, bankruptcy costs outweigh transactions costs. In response, firms optimally reduce debt maturity. Conversely, a flat or negative term spread reduces the risk of financial distress and in response, firms optimally increase their debt maturity. In other words, as the economy cycles through, the resulting term structure creates natural incentives for firms to shift their debt maturity.

In our framework, firms choose debt maturity to maximize the total firm value. Specifically, there are no agency problems between debt holders and equity holders. Firm's managers and debt investors are perfectly informed about all the relevant variables that characterize the firm and the economy. Furthermore, there is no liquidity risk in our setup. At time zero, the firm issues a T-year coupon bond after paying transaction costs related to debt issuance. If the firm has not gone bankrupt in T years, the firm rolls over its debt by issuing a new T-year coupon bond at time T, after paying transaction costs. If at the end of the second T-year period the firm is still solvent, it issues another T-year coupon bond. This process goes on indefinitely as long as the firm is solvent. If the firm goes bankrupt, debt holders take over the firm's operations after paying bankruptcy costs.

We extend the existing modeling literature by incorporating the additional stylized facts:

1. **Target leverage ratio** — According to the survey results of Graham and Harvey (2001), 44% of CFOs report to have a strict or somewhat strict target leverage ratio. 37% claim to have a flexible target leverage ratio. Remarkably, only 19% of the CFOs claim that they have neither a target ratio nor a target range. Lemmon, Roberts, and Zender (2008) reinforce the evidence of target leverage ratios. They show that capital structures are remarkably stable over time; and that firms with high (low) leverage maintain relatively high (low) leverage for over twenty years, independent of being public or private. Frank and Goyal (2003), Leary and Roberts (2005), Flannery and Rangan (2006), and Huang and Ritter (2009) also corroborate that managers adjust their capital structure towards a specific target.

2. **Lumpy debt maturity** — Choi, Hackbarth, and Zechner (2011) provide empirical evidence confirming that the debt maturity structure of firms is lumpy and not granular. That is, debt maturity tends to be concentrated as opposed to being scattered across different points in time. For example, if debt maturities are distributed uniformly in an interval $[\underline{T}, \overline{T}]$, then the maturity structure is granular. In our setup, firms issue debt with maturity T at every rollover date, characterizing a lumpy maturity structure.

The short rate follows a mean-reverting process as in Vasicek (1977). This approach allows us to examine the impact of the dynamics of the term spread. The term spread is a state variable that proxies economic conditions. As the term spread cycles, so does the economy. In the risk neutral measure, the drift of the firm value is the short rate. Therefore, when the

term spread is relatively high, i.e. when the short rate is significantly lower than the long rate, the firm growth rate is initially low and it accelerates towards its long run equilibrium. This indicates that economic recovery is ahead as would be the case at the end of a recession. On the other hand, when the term spread is relatively low, the firm growth rate is initially high and it decelerates towards its long run equilibrium. This indicates grim economic future as would be the case at the beginning of a recession.

The economic intuition about the link between term spread and debt maturity is the following. When the term spread is significantly positive, the short rate is significantly below the long run equilibrium rate. Even though the short rate increases on average in the future, it is expected to remain below the long rate for a significant time. With the prospect of low growth ahead, the firm's probability of default is high. Therefore, the firm chooses to decrease maturity at the expense of paying higher transaction costs related to debt rollover. Conversely, when the term spread is negative, the short rate is greater than the long rate. In this case, the probability of default is low because the firm is expected to grow at higher than normal rates. Therefore, the firm chooses to increase maturity to minimize the transaction costs related to debt rollover.

Our paper contributes to the existing literature in three different ways. First, it adds to the corporate finance literature that deals with the speed of adjustment in the firm's capital structure. There is a debate in the literature about how frequently firms adjust their capital structure. Strebulaev (2007) employs a tradeoff theory model to show that firms adjust their capital structure infrequently. That is, managers optimally choose to be inactive in the process of maximizing firm value. Conversely, Welch (2012) uses empirical evidence to argue that managers are actually quite active in adjusting their capital structure. Our results indicate that the optimal maturity of firms varies between 1-3 years for a wide range of parameters. This level of activity is in line with Welch (2012). In common to ours and Strebulaev (2007)'s framework, firms incur transaction costs during debt rollover. The main difference, however, concerns the choice of debt. In his model, firms either increase or decrease leverage by issuing perpetual debt. In our model, firms issue a finite T-maturity debt to reach a target leverage ratio. Consequently, firms are forced to readjust their capital structure every T periods in our model while small transaction costs lead to large waiting time (inactivity) in his model.

Second, our paper adds to the literature concerning debt maturity and systematic risk, as measured in our model by the term spread. Chen, Xu, and Yang (2012) also explain this link in a setting with liquidity risk, whereas in our framework, markets are perfectly liquid. Lastly, our paper is closely related to Ju and Ou-Yang (2006) who explain the effect of stochastic interest rates on leverage, debt maturity, and credit spreads. To achieve a closed form solution, Ju and Ou-Yang (2006) assume that debt is not issued at par during debt rollover. An artifact of this assumption is that debt maturity is independent of the short rate, which is inconsistent with the empirical evidence. We enhance their model by

assuming that debt is issued at par and that firms adjust their capital structure toward a target leverage ratio. These features allow our model to produce results that are consistent with the empirical data, showing that debt maturity is inversely related to the term spread.

Our theoretical predictions match the empirical findings of Barclay and Smith (1995) and Julio, Kim, and Weisbach (2008). The first prediction concerns leverage. Julio, Kim, and Weisbach (2008) find that there is no substantial difference in leverage between firms that issue short-term and long-term debt. In our model, this result is mechanical as the firms re-balance their capital structure towards a target leverage ratio. The second prediction concerns volatility. Barclay and Smith (1995) and Julio, Kim, and Weisbach (2008) find that more volatile firms issue shorter term debt. They attribute this empirical finding as a validation of the agency theory of Myers (1977). Our setup matches this finding naturally. When the firm value is more volatile, big changes in the firm value are more likely. Therefore, the firm optimally decides to re-balance its capital structure more frequently as its probability of default is higher. To summarize, we find that the long run interest rate, the volatility of the interest rate process, the correlation between the short rate and the firm value, and the volatility of the firm's value are all important parameters for determining the optimal debt maturity.

The remainder of this article is organized as follows. Section II presents motivating empirical evidence that shows the inverse relationship between debt maturity and term spread. Section III presents the model setup. In order to get an intuition for the link between interest rates and default risk, we analyze the value of a risky zero coupon bond in Section IV. Section V derives the levered firm value using the trade-off theory of capital structure. Section VI presents the quantitative analysis. Section VII summarizes the article and makes concluding remarks.

2.2 Empirical evidence between debt maturity and term spread

We divide this section into two parts. First, we analyze the time series data of aggregate debt maturity using data from the Flow of Funds Accounts. Specifically, this part shows that on the aggregate, firms reduce debt maturity when the term spread is significantly positive and they increase maturity as the term spread decreases. We also show that aggregate debt maturity is high prior to the beginning of a recession and it decreases by the end of a recession. Second, we review cross-sectional evidence of debt maturity and term spread from the corporate finance literature.

Analysis of aggregate debt maturity and term spread

The dark solid line in Figure 2.1 shows the cyclical component of the share of the long term debt[1] calculated by applying the Hodrick-Prescott filter[2]. The data pertains to debt issued by non-financial corporate firms. The maturity is classified as "long term" if it is greater than one year. The shaded bands indicate recessions as designated by NBER. The link between the cyclical share of long term debt and the macroeconomic conditions is clear. During recessions, the long term debt share appears to dip below the trend. For example, during the first quarter of 2008, at the start of the past recession, the share of the long term debt was 3% below the trend. During the third quarter of 2009, the first quarter after the end of the past recession, the share of the long term debt was only 0.04% below the trend.
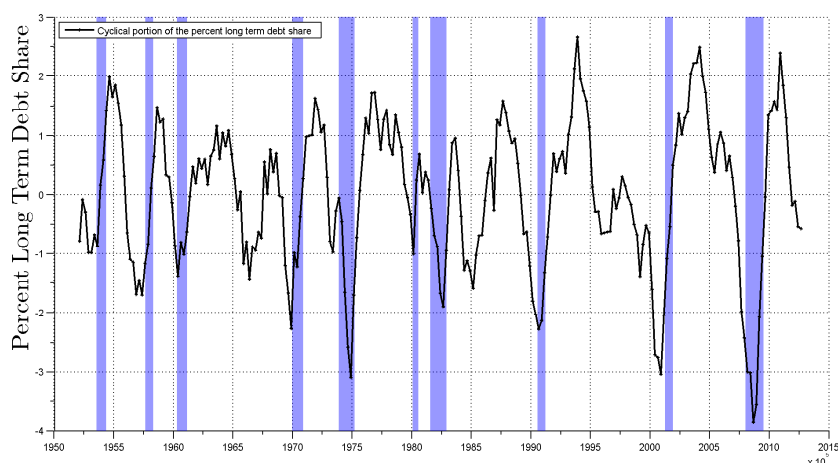
Figure 2.1: This is a time series plot of the long term debt share of non-financial corporate business. The dark line is the cyclical component of the long term debt share calculated via the Hodrick-Prescott filter. The data for the long term debt share is from the Fed Funds flow database (series L.102). The shaded bands in gray are the NBER recession dates.

Table 2.1 contains descriptive statistics of the share of long term debt since the first quarter of 1952. The share of the long term debt is 1.31% below the trend during recessions, while the share is 0.23% above the trend during non-recessionary times. Chen, Xu, and Yang (2012) and Julio, Kim, and Weisbach (2008) corroborate the results above by showing that debt maturity decreases during recessions.

We now focus on the relationship between term spread and the state of the economy. The dark solid line in Figure 2.2 shows the cyclical portion of the term spread. A negative term spread is often a harbinger of a recession. Estrella and Hardouvelis (1991) shows that there

---

[1]The data for the long term debt share is from the Fed Funds flow database (series L.102). The data spans from first quarter of 1952 to third quarter of 2012 — a total of 243 quarters. Additionally, the data spans 10 recessions — a total of 36 quarters.

[2]For completeness, we plot the trend portion of the share of the long term debt in subsection 5.1 of the Appendix.
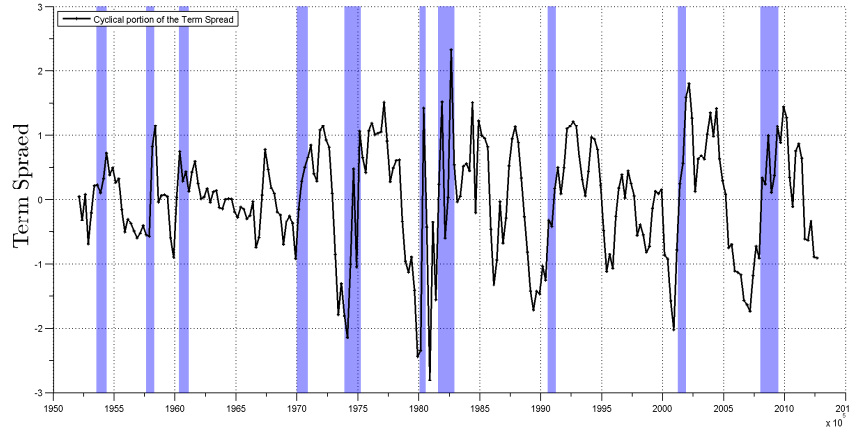
Figure 2.2: This is a time series plot of the cyclical portion of the term Spread (difference between 10-year Treasure note yield and the 3-month Treasury bill). The data is taken from Global Financial database. The shaded bands in gray are the NBER recession dates.

Event	Mean	Std. Dev	Min	Max
Recession	-1.31	1.09	0.57	-3.86
No Recession	0.23	1.10	2.67	-3.03

Table 2.1: Descriptve statistics of the percent long term debt share

is a marked dip in the term spread about six quarters prior to a recession. Figure 2.2, for instance, shows that the term spread first became negative in the third quarter of 2006, approximately five quarters before the beginning of the past recession. The dynamics of the term spread are even more evident from the descriptive statistics in Table 2.2. Three quarters prior to a recession, the term spread is flat, averaging 0.05% over the past ten recessions. Then, it increases to an average of 0.93% at the beginning of a recession. It increases to its apex of 2.30% by the end of the recession. Finally, three quarters after a recession, the term spread begins to decrease, reaching an average of 2.10%.

Event	Mean	Std. Dev
Term spread three quarters prior to a recession	0.05%	0.72%
Term spread at the beginning of a recession	0.93%	1.21%
Term spread at the end of a recession	2.30%	0.86%
Term spread three quarters after the end of a recession	2.10%	0.95%

Table 2.2: Dynamics of the term spread

From both figures and descriptive statistics tables, it is evident that debt maturity and term spread are intimately linked. We now review cross-sectional evidence from the corporate finance literature.

Cross-sectional evidence linking debt maturity and the term spread

Empirical evidence finds that managers time their borrowing activity. They use their beliefs about future interest rate movements to lower the cost of funds. A prominent example is the CFO survey conducted by Graham and Harvey (2001). Based on their results, CFOs' say that they issue short-term debt when "short-term rates are low compared to long-term rates," or when "we are waiting for long-term rates to come down."

The fact that managers are speculating is further reinforced by Faulkender (2005)'s empirical study of financial policies for firms in the chemical industry between 1994 and 1999. Faulkender finds that firms that issue floating rate debt do not swap floating interest payments for fixed interest payments. Firms seem to be amplifying their interest rate risk as opposed to reducing their interest rate risk.

Using data from corporate bond issuances, Barclay and Smith (1995), Guedes and Opler (1996), Faulkender and Petersen (2006) and Julio, Kim, and Weisbach (2008) also find that firms issue shorter-term debt when the term spread is significantly positive. This result is robust to the addition of other firm specific variables such as credit ratings, book to market ratio, and stock return volatility.

In our paper, we provide a natural explanation of why firms vary debt maturity with term spread. Based on our model, as the economy cycles through, the resulting term structure creates natural incentives for firms to shift their debt maturity. The explanation does not rely on either agency conflicts, or asymmetric information. This is also consistent with Graham and Harvey (2001), who show evidence that the CFOs are not concerned with either agency conflicts or asymmetric information, when issuing debt. Finally, they show that CFOs are concerned with transaction costs associated with debt rollover, which supports our theoretical approach.

2.3   Model Setup

In this section, we modify the setup of Leland and Toft (1996) in three ways. First, we relax the assumption of constant interest rates by assuming that the short rate process follows a mean reverting process. This assumption implies that the term spread is stochastic which in turn allows us to link debt maturity and term spread.

Second, for expositional clarity, we assume exogenous default boundary as opposed to endogenous default boundary. The notion of exogenous versus endogenous boundary is important when one is concerned with agency conflicts. For example, the debt holder's incentives to default are obviously different from the equity holder's incentives. We show in Section IV that the choice of exogenous default boundary produces the same qualitative results for credit spreads as other structural credit models with endogenous default boundary.

Third, we assume that our default boundary, which is in the spirit of Black and Cox (1976), is stochastic and not constant. This approach may seem counterproductive, but we show that our parametric form of default boundary allows us to evaluate the model in closed form. Particularly, with our parametric form, we show that distance to default is directly related to the short rate. That is, when the short rate is high, firms are less likely to default. This feature forms the basis of our model.

We perform our analysis in partial equilibrium and we assume complete markets. This allows us to perform our analysis directly in the risk-neutral measure. The details of the assumptions are discussed below.

Environment

ASSUMPTION 4. (Interest rate dynamics) *Let r_t denote the short-term riskless interest rate. The dynamics of r_t are given by*

$$dr_t = \beta\left(\alpha - r_t\right)dt + \sigma_r\,d\mathbb{W}_{rt}, \tag{2.1}$$

where β, α and σ_r are constants and $\mathbb{W}_{rt}$ is a standard Wiener process in the risk-neutral measure.

The dynamics of r_t is drawn from the term structure model of Vasicek (1977). Let $\Lambda(r_t, t, T)$ be the time t price of a zero coupon bond with maturity T. Standard calculations yield the following expression for the value of a riskless bond:

$$\Lambda(r_t, t, T) = e^{A(t,T) - B(t,T)\,r_t}, \tag{2.2}$$

where

$$B(t,T) = \frac{1 - e^{\beta(T-t)}}{\beta}; \quad A(t,T) = (\alpha - \frac{\sigma_r^2}{2\beta})\left[B(t;T) - (T-t)\right] - \frac{\sigma_r^2 B(t;T)^2}{4\beta}.$$

We define the term spread as the difference in the yield of a ten year zero coupon Treasury bond and a 3-month zero coupon Treasury bill. Mathematically, the term spread

$TermSpread(r_0, T)$ is

$$TermSpread(r_0, T) \ = \ -\frac{\ln \Lambda(r_0, 0, 10)}{10} \ - \ (-\frac{\ln \Lambda(r_0, 0, 0.25)}{0.25}).$$

Loosely speaking, the term spread is well approximated as

$$TermSpread(r_0, T) \ \approx \ \alpha \ - \ r_0.$$

ASSUMPTION 5. (Firm dynamics) *Let V_t designate the market value of the firm's unlevered assets before tax. The dynamics of V_t are given by*

$$\frac{dV_t}{V_t} = (r_t - y)\, dt \ + \ \sigma_V\, d\mathbb{W}_{vt}, \tag{2.3}$$

where y and σ_V are constants, $\mathbb{W}_{vt}$ is also a standard Wiener process. The instant correlation between $d\mathbb{W}_{vt}$ and $d\mathbb{W}_{rt}$ is $\rho\, dt$[3].

The firm pays a constant fraction y of its unlevered assets to its equity holders as dividends. If the firm does not issue any debt, equity holders are entitled to a fraction $(1 - \theta)$ of the dividends generated from the firm's unlevered assets. In this case the after-tax unlevered value of the firm is $(1 - \theta)V_t$ where θ is its marginal corporate tax rate.

ASSUMPTION 6. (Default threshold dynamics) *Following Black and Cox (1976), we assume there is an exogenous threshold value V_D at which the firm defaults on its debt. The threshold $V_D(t)$ is given by*

$$V_D(r_t, t, T) = P\Lambda(r_t, t, T)\, e^{y(T-t)} \, / \, (1 - \theta) \tag{2.4}$$

where P is the face value (principal) of debt. The debt holder receives $1 - \gamma$ times the after-tax unlevered firm value upon default.

Note that the threshold has the desired property that at the maturity date T, its value is equal to the face value of debt. Therefore, as long as the after-tax market value of assets $(1 - \theta)\, V_T$ remains higher than $(1 - \theta)\, V_D$, debt holders will be compensated fully. At any time prior to T, the threshold value indicates that the after-tax market value of assets should remain higher than the present value of the principal. This expression for the threshold is in the spirit of that used by Black and Cox (1976) after adjusting for stochastic interest rates and dividend yield. This definition of financial distress is consistent with covenants referring

---

[3]Technically, uncertainty is described by two dependent Brownian motions, $\{\mathbb{W}_{vt}, \mathbb{W}_{rt}\}$ for $t \geq 0$ defined on a complete probability space $(\Omega, \mathbf{F}, \mathbb{Q})$ where $\mathbf{F} = \mathcal{F}_{\{t \geq 0\}}$ is the augmented filtration generated by $\{\mathbb{W}_{vt}, \mathbb{W}_{rt}\}$.

to a violation of minimum net worth or working-capital requirements as implemented in Kim, Ramaswamy, and Sundaresan, 1993.

Most notably, the threshold value is exogenous and not endogenous as in Leland (1994) or Leland and Toft (1996). The magnitude of the difference in credit spreads between assuming exogenous versus endogenous bankruptcy boundary is not significant as highlighted in Fong (2006).

Note that the default threshold V_D depends on r_t and hence it is stochastic. At first glance, imposing such structure might seem counterproductive. However, the study of first passage time of default is made simpler by studying the dynamics of

$$X_t \equiv \log \left[\frac{V_t}{V_{D_t}} \right],$$

which measures distance to default. By Ito's Lemma, the dynamics of X_t are

$$dX_t = \left[\frac{\sigma_p^2(t;T)}{2} - \frac{\sigma_V^2}{2} \right] dt + \sigma_V d\mathbb{W}_{vt} + \sigma_p(t,T) d\mathbb{W}_{rt}. \tag{2.5}$$

where $\sigma_p(t,T) \equiv \sigma_r B(t;T)$. Note that both the drift and volatility of X_t simplify to a deterministic expression. This is crucial for obtaining a closed form expression of the first passage time until default, as shown in a later section.

Figure 2.3 shows the relationship between the distance to default measured by X and the short rate r_0 for different times to maturity T. For a given time to maturity, the distance to default X is always increasing function of the short rate. This is because the higher short rate, the higher the drift of the firm value. Therefore, for a given bond with principal P_0, there is a lower chance that the firm value will breach the default threshold. The sensitivity of the distance to default X and short rate r_0 increases with time to maturity. This is evident from the graph in which the dotted dashed line representing a maturity of 3 years has a higher slope than the solid line representing a maturity of 1 year.

The fact that X increases with short rate is crucial for our results. A high short rate means that the term spread is low. Consider the decision making process of a firm when the term spread is low. Suppose the manager chooses a maturity T_1 say 1 year. From the Figure 2.3, it is clear that the distance to default measure is high which means that probability of bankruptcy is low. In this case, the manager can afford to increase maturity to $T_2 > T_1$ to minimize the transaction costs associated with debt rollover. Conversely, if the short rate is low so that the term spread is high, the probability of default is high. Therefore, the manager will optimally decrease maturity to minimize bankruptcy costs.

ASSUMPTION 7. (Debt rollover dynamics) *The firm adjusts its capital structure every T years. At time zero, the firm issues a T-year coupon bond with principal P_0 and*
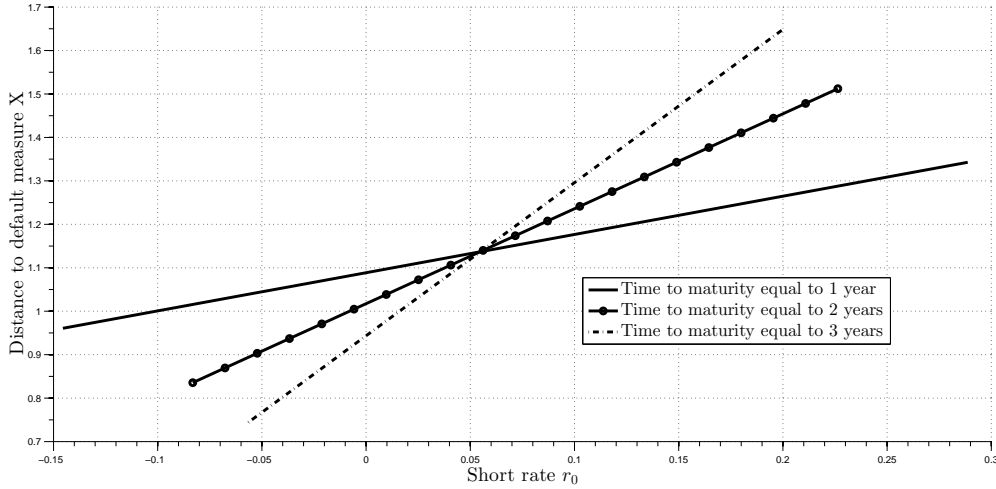
Figure 2.3: This figure shows the relationship between the distance to measure X and the short rate r_0 for different time to maturity T.

coupon rate c_0. *The firm chooses principal P_0 to achieve an exogenously specified target leverage ratio. If the firm does not default in T years, it issues another T-year coupon bond at time T. This process continues indefinitely as long as the firm is solvent. The firm incurs a transaction cost of ϕ times the value of debt at every debt issuance.*

Most models of trade-off theory assume either (i) perpetual debt (model based on Leland (1994))[4] or (ii) static debt (model based on Leland and Toft (1996)). However, these types of specifications are clearly not suitable for analyzing debt optimal maturity. The assumption of finite maturity debt is critical for analyzing optimal debt maturity.

ASSUMPTION 8. *(Market value of debt) Debt is issued at par. Specifically, at every debt issuance nT where $n \in \{0, 1, 2, ...\}$, conditional upon not defaulting prior to nT, the market value of debt, L_{nT}, is equal to the face value of the bond, i.e. $L_{nT} = P_{nT}$.*

Assumption 8 is standard in the trade-off theory literature.

ASSUMPTION 9. *(Specific target leverage ratio) At every debt issuance nT where $n \in \{0, 1, 2, ...\}$, conditional upon not defaulting prior to nT, the manager adjusts the capital structure toward a specific target leverage ratio ζ, where*

$$\zeta \equiv \frac{P_n}{(1 - \theta) V_{nT}} = \frac{P_0}{(1 - \theta) V_0}.$$

---

[4]Model that allow for capital structure adjustment also assume perpetual debt as in Goldstein, Ju, and Leland (2001) and Strebulaev (2007)

This assumption is in the spirit of the stylized facts identified by Graham and Harvey (2001) and Lemmon, Roberts, and Zender (2008), which states that managers adjust their capital structure toward a specific target leverage ratio. Note that in our setup, leverage is not a choice variable. That is, firms only choose maturity to maximize the firm value. In this manner, we differ from the existing literature on trade-off theory, which focuses primarily on maximizing firm value by choosing the leverage ratio.

In order to better understand the implications of Assumption 9, it is useful to show the firm, debt issuance and default dynamics with a hypothetical sample path as in Figure 2.4. In this example, we set the time to maturity T to four years and the target leverage ratio to 64%. Therefore, at time zero, the firm issues a bond with a face value of \$42.00. The black line shows the sample path of the unlevered firm value $\{V_t\}$ prior to default. The red line depicts the default threshold $\{V_D\}_t$. From equation (2.4), the dynamics of the default threshold depend on the dynamics of the interest rate $\{r_t\}$. Since the volatility of the interest rates is low, the fluctuations in the dynamics of the default threshold are smaller than the fluctuations in the firm value.

Note that the firm value remains above the default threshold for the first four years. In this first debt issuance, the firm does not default and it readjusts its capital structure to the specified target leverage ratio. Furthermore, since the asset value increases at the end of the fourth year, the firm rolls its debt by issuing a bond with a higher face value of \$80.70. A little before year six, the firm value breaches the default threshold. At this point, the firm declares bankruptcy, and debt holders lose a fraction γ of the after-tax unlevered firm value due to bankruptcy costs. The gray line shows the hypothetical firm value after debt holders take over the operations of the firm. To summarize, our model set up is a repeated version of Leland and Toft (1996) with the following modifications: stochastic interest rates, exogenous and stochastic default threshold, and exogenous target leverage ratio. Prior to analyzing the tradeoff between the tax benefits and the sum of bankruptcy costs and transaction costs related to debt rollover, it is useful to analyze the value of a simpler security — a zero coupon risky bond. We use the change of numeraire technique to derive a closed-form expression for the value of such security. This technique will be used again in later section to derive the optimal debt maturity.

2.4   Valuation of a risky zero coupon bond

In this section, we value a hypothetical zero coupon risky bond to show three features of the model. First, and most importantly, we show that the credit spreads are directly related to the term spread. That is, when the term spread is low, so that the short rate is high, the probability of default is low. Therefore, the firm can afford to increase the maturity to minimize transaction costs of debt rollover. Second, we show that our model can reproduce credit spreads that are broadly consistent with other structural models of capital structure.
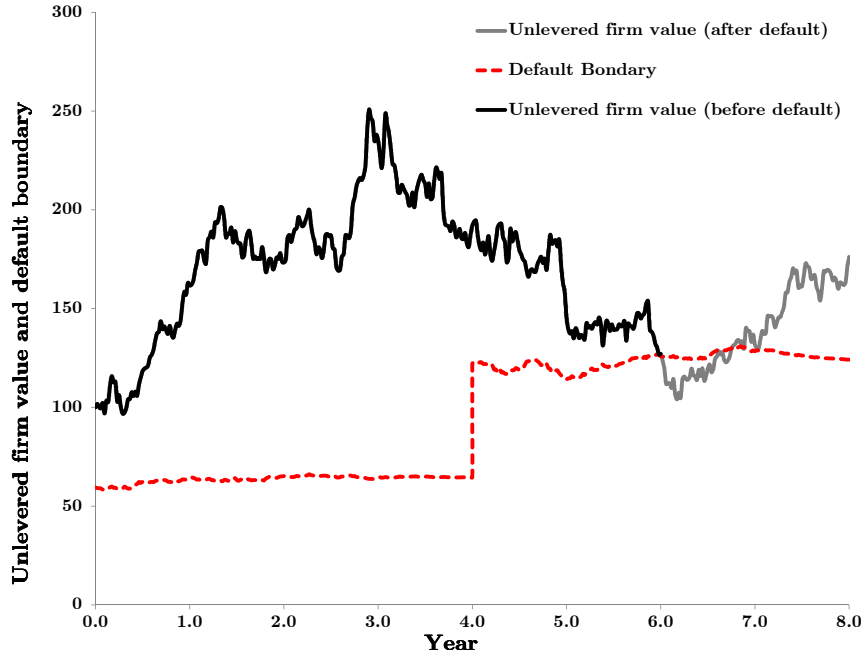
Figure 2.4: This figure shows a sample path of the firm value and default threshold. The simulation shows a sample path in which 1) the firm does not default at the end of the 4th year (the time to maturity), 2) it defaults after it re-adjusts its capital structure in the 6th year. Furthermore, the firm rebalances, so that the log ratio $\ln \frac{V_0}{P_0}$ is the same at every re-adjustment date. The dark black and gray lines show the unlevered firm values. The gray line shows the entire sample path of the firm value even though firm defaults in the 6th year. The red line shows the stochastic default threshold.

The third motivation of this section is technical in nature. We show that the technique of change of numeraire allows us to get a closed form expressions for the value of a risky zero coupon bond.

Setup

Let $D_{\text{zero}}(t, T, r_t; X_0)$ denote the price of a risky zero coupon bond with maturity date T at time $t \leq T$. The payoff on this contingent claim is \$1 if default does not occur during the life of the bond, and \$$(1 - \gamma)$ otherwise. This payoff function is expressed as

$$1 - \gamma \, \mathbb{I}(\text{Default happens prior to } T).$$

where $\mathbb{I}$ is an indicator function that takes the value one if V_t reaches V_{D_t} during the life of the bond, and zero otherwise. Since both V_t and V_{D_t} are stochastic, it is prudent to work with their ratio. From the definition of $X_t = \ln \frac{V_t}{V_{D_t}}$, default takes place when X_t reaches zero from above. Formally, $\mathbb{I}$ takes the value of one if $\tau \leq T$ where

$$\tau \doteq \inf\{t \geq 0 : X_t \leq 0\},$$

or zero otherwise. Therefore, the value of the risky zero bond is

$$
\begin{aligned}
D_{\text{zero}}(t, T, r_t; X_0) &= \mathbb{E}_t \left[e^{\int_t^T r_u \, du} \times \{1 - \gamma \, \mathbb{I}(\tau \leq T)\} \right] \\
&= \mathbb{E}_t \left[e^{\int_t^T r_u \, du} \times 1 \right] - \gamma \, \mathbb{E}_t \left[e^{\int_t^T r_u \, du} \times \mathbb{I}(\tau \leq T) \right] \qquad (2.6)
\end{aligned}
$$

The first term represents the present value of one dollar upon no default. This expression is simply the value of a default free zero coupon bond $\Lambda(r_t, t, T)$. The second term represents the loss given default. The expectation depends on the sample path of both interest rates and firm values and hence it is difficult to derive in closed-form upon first glance. In the fortunate case when the interest rate process and firm values are independent, the second term can be written as a product of two expectations, i.e.

$$
\begin{aligned}
\mathbb{E}_t \left[e^{\int_t^T r_u \, du} \times \mathbb{I}(\tau \leq T) \right] &= \mathbb{E}_t \left[e^{\int_t^T r_u \, du} \right] \times \mathbb{E}_t \left[\mathbb{I}(\tau \leq T) \right] \\
&= \Lambda(r_t, t, T) \times \Pr(\tau \leq T).
\end{aligned}
$$

Given the dynamics of X_t in equation 2.5, the first passage time probability can be calculated in closed form by using the Kolmogorov backward equation. However, it is hard to justify independence between firm value and interest rate process in the risk-neutral measure[5]. The change of measure technique, which is introduced in the Appendix (subsection 5.3), allows us to write the second term on the right hand side in equation 2.6 as a product of two expectations without assuming independence.

In subsection 5.3 of the Appendix, we show that the

$$
\mathbb{E}_t \left[e^{\int_t^T r_u \, du} \times \mathbb{I}(\tau \leq T) \right] = \Lambda(r_t, t, T) \times \mathbb{E}_t^{\mathbb{T}} \left[\mathbb{I}(\tau \leq T) \right] = \Lambda(r_t, t, T) \times \Pr^{\mathbb{T}}(\tau \leq T),
$$

where $\mathbb{E}_t^{\mathbb{T}}$ is calculated using a new measure $\mathbb{Q}^{\mathbb{T}}$. Mathematically, the expectation above is the cumulative distribution function of the first passage time evaluated in the new $\mathbb{Q}^{\mathbb{T}}$ measure. The firm value as normalized by the price of a default free bond is a martingale in the new measure. In other words, in the $\mathbb{Q}^{\mathbb{T}}$ measure, a T maturity default free zero coupon bond is used as the numeraire. By comparison, the money market account is used as the numeraire in the risk neutral measure. After deriving the dynamics of $\{X_t\}$ in the new $\mathbb{Q}^{\mathbb{T}}$ measure, we express the distribution of the first passage time in closed form using the Kolmogorov backward equation.

To summarize, the value of a zero coupon bond at time $t = 0$ is

$$
D_{\text{zero}}(0, T, r_0; X_0) = \Lambda(r_0, 0, T)(1 - \gamma \, G(T, T, X_0)),
$$

where

$$
G(t, T, X_0) = \frac{1}{\Lambda(r_0, 0, T)} \mathbb{E}_t \left[e^{-\int_0^T r_u \, du} \times \mathbb{I}(\tau \leq t) \right].
$$

The exact expression for $G(.)$ is given in equation (5.5) in the Appendix.

---

[5]In the risk neutral measure, the drift of V_t is the $r_t - y$ and hence they cannot be independent.

Credit spread and the shape of the term structure

Given the explicit solution for risky zero coupon bond, we can solve for the credit spread, which is defined as the difference between the yields of a risky and a riskless bond with same maturity. Figure 2.5 graphs the term structure of credit spreads for a low leveraged firm. For this example, we set the principal P_0 to $ 20.00 and the after tax firm value to $65.00 $=$ $100.00 \times (1 - 0.35)$. The figure graphs the credit spreads for different levels of the short rate r_0. We choose the parameters α, β, and σ_r governing the short rate stochastic process to closely match the observed moments given in Ju and Ou-Yang (2006). The term structure of credit spreads are monotonically increasing as a function of maturity. This result aligns well with the empirical evidence found by Sarig and Warga (1989), who suggest that the term structure of credit spreads increases with maturity for bonds with high credit ratings.

Figure 2.6 graphs the term structure of credit spreads for highly leveraged firms where we set the principal P_0 to $50.00, while keeping the same value for the other parameters. The term structure of credit spreads is hump shaped. This is also consistent with Sarig and Warga (1989).

The credit spread is directly proportional to the term spread. From both graphs, the credit spreads are higher when the term spread is higher. This pattern is true for both highly leveraged and low levered firms. Lastly, the concavity of the term structure of credit spreads for intermediate maturities is also dependent on the term spread. The concavity is directly proportional to the term spread.

We conclude this section by pointing out that our default mechanism generates credit spreads consistent with other well established structural default models. In the next section, we develop a model in which firms optimally choose maturity to maximize the firm value.

2.5 Levered firm value

In this section, we derive the firm value taking into account the tradeoff between tax benefits and the sum of bankruptcy costs and transaction costs. We divide this section into two parts. In the first part, we derive closed form expressions for the present value of tax benefits, bankruptcy costs and transaction costs for one bond issuance. In the second part, using a fixed point argument, we solve for the firm value considering infinite debt issuances.

One time debt issuance

Assume that the firm issues a T maturity bond so that the leverage ratio is equal to the target level ζ. This means that the firm issues a bond with principal P_0 to satisfy Assumption 9.

Figure 2.5: This figure shows the credit spreads of low levered firms for different shapes of the term structures. The parameters are as follows: $\beta = 0.261$, $V_0 = 100$, $\alpha = 0.0716$, $\gamma = 0.5$, $\theta = 0.35$, $y = 0.05$, $P_0 = 50$ and $r_0 = 0.01, 0.07, 0.13$ for upward sloping, flat and downward sloping term structures. Credit spread is the implied yield of the bond minus the short rate.

Figure 2.6: This figure shows the credit spreads of highly levered firms for different shapes of the term structures. The parameters are as follows: $\beta = 0.261$, $V_0 = 100$, $\alpha = 0.0716$, $\gamma = 0.5$, $\theta = 0.35$, $y = 0.05$, $P_0 = 20$ and $r_0 = 0.01, 0.07, 0.13$ for upward sloping, flat and downward sloping term structures. Credit spread is the implied yield of the bond minus the short rate.

Expression for transaction cost

From Assumptions 7 and 8, the transaction cost of issuing debt is given by

$$tc(r_0, \zeta, T, V_0) \ = \ \phi \times L_0 \ = \ \phi \times P_0 \ = \ V_0 \times ntc(\zeta) \tag{2.7}$$

where

$$ntc(\zeta) \equiv \phi \times \zeta \times (1 - \theta).$$

The function $ntc(\zeta)$ can be interpreted as transaction costs per unit of unlevered firm value, so it is the normalized transaction cost. Equation (2.7) shows that the transaction cost is a function of the target leverage ratio and more importantly, the transaction cost is a linear function of the unlevered firm value V_0.

Expression for bankruptcy cost

When default occurs, a fraction γ of the unlevered firm value is lost in bankruptcy procedures. Formally, bankruptcy cost is

$$bc(r_0, \zeta, T, V_0) = \mathbb{E}_0 \left[\int_0^T ds\, \gamma\, V_D(r_s, s, T)\, \delta(s - \tau)\, e^{-\int_0^\tau r_u du} \right],$$

where $\delta(.)$ is the dirac-delta function. Consider the term inside the integral. Suppose default takes place at some time $\tau = s \in [0, T]$ so that the function δ takes a value of 1 at that moment. The bond holders will recover the present value of γV_τ. At default, the firm value is equal to the default threshold, i.e. $V_{\tau=s} = V_D(r_s, s, T)$. Therefore, the term in the integral is the present value of the loss suffered by the bondholders in the event default takes place at time τ. The integral represents the loss considering the fact that default can take place at any time between 0 and T. In the Appendix, we show that the expression for the bankruptcy cost is

$$bc(r_0, \zeta, T, V_0) = V_0 \times nbc(r_0, \zeta, T) \tag{2.8}$$

where

$$nbc(r_0, \zeta, T) \ = \ \zeta\, \gamma\, \Lambda(r_0, 0, T) \left[G(T; T, X_0) + \widehat{G}(T; T, X_0) \right],$$

and

$$\widehat{G}(T; T, X_0) = y \int_0^T ds\, e^{y(T-s)}\, G(s; T, X_0).$$

The function $nbc(r_0, \zeta, T)$ can be interpreted as the bankruptcy cost per unit of unlevered firm value or the normalized bankruptcy cost. Note that bankruptcy cost is a linear function of the unlevered firm value V_0. Also note that X_0 is implicitly a function of the short rate r_0.

Expression for tax benefit of debt

When the coupon of C is paid, a fraction θC is deducted from corporate taxes. Furthermore, the firm only enjoy tax benefits if it remains solvent. Formally, the tax benefit for one debt issuance is

$$tb(r_0, \zeta, T, V_0) = \mathbb{E}_0 \left[\int_0^T ds \, \mathbb{I}_{s < \tau} \, \theta C \, e^{-\int_0^s r_u du} \right].$$

The integral represents the present value of tax benefit considering the fact that default can take place at any time between 0 and T. Suppose default takes place at some time $\tau \leq T$ so that $\mathbb{I}$ takes a value of one in the interval $[0, \tau]$ and zero otherwise. The interval $[0, \tau]$ represents the times in which the firm was solvent.

We can evaluate the expression for tax benefit indirectly by using Assumption 8. The market value of debt is given by

$$
\begin{aligned}
L_0 &= \mathbb{E}_0 \left[\int_0^T C e^{-\int_0^s r(u) du} \mathbb{I}_{s < \tau} ds \right] \\
&+ \mathbb{E}_0 \left[P_0 \mathbb{I}_{\tau > T} e^{-\int_0^T r_u du} \right] \\
&+ \mathbb{E}_0 \left[(1 - \theta)(1 - \gamma) \int_0^T V_D(r_s, s, T) \delta(s - \tau) e^{-\int_0^\tau r_u du} ds \right].
\end{aligned}
$$

The value of debt is composed of three parts: (i) present value of the flow of coupon payments prior to maturity while the firm remains solvent; (ii) present value of principal payment at time T conditional upon not defaulting prior to T and (iii) the present value of the recovery amount conditional upon defaulting at any time before T.

In subsection 5.5 of the Appendix, we show that the expression for tax benefit is

$$tb(r_0, \zeta, T, V_0) = V_0 \times ntb(r_0, \zeta, T) \tag{2.9}$$

where

$$ntb(r_0, \zeta, T) \equiv \theta(1-\theta)\zeta \left[1 - \Lambda(r_0, 0, T)(1 - G(T, T, X_0)) - (1 - \gamma) \{ G(T, T, X_0) + \widehat{G}(T, T, X_0) \} \right].$$

The function $ntb(r_0, \zeta, T)$ can be interpreted as the tax benefit per unit of unlevered firm value, so it is the normalized tax benefit. Note that tax benefit is a linear function of the unlevered firm value V_0.

At this point, we can also back solve for the value of the coupon rate C. Mathematically,

$$C = \frac{tb(r_0, \zeta, T, V_0)}{\theta \, \mathbb{E}_0 \left[\int_0^T e^{-\int_0^s r(u) du} \, \mathbb{I}_{s < \tau} \, ds \right]} = \frac{tb(r_0, \zeta, T, V_0)}{\theta \, \tilde{G}(T, T, X_0)}, \tag{2.10}$$

where

$$\tilde{G}(T, T, X_0) \; = \; \int_0^T ds\, \Lambda(r_0, 0, s)\,(1 - G(s, s, X_0)).$$

The details of the derivation are given in subsection 5.6 of the Appendix.

Infinite debt issuances

Markov-Chain approximation of $\{r_t\}$

In order analyze the firm value with infinite debt issuances, it is useful to approximate the continuous time interest rate process $\{r_t\}$ by a Markov-Chain. It is well-known that the Vasicek interest rate process can be expressed as an AR(1) process. We closely follow the approach outlined in Tauchen (1986), who discusses accuracy of approximating an AR(1) process with a Markov-Chain.

First note that the conditional expectation and variance of the short rate process in equation (2.1) are given by

$$r_{s|t}^{\mathrm{mean}} \; \equiv \; \mathbb{E}_t\left[r_s\right] \; = \; \alpha \; + \; (r_t - \alpha)\, e^{-\beta(s-t)} \qquad \text{for } s \geq t;$$

and

$$r_{s|t}^{\mathrm{var}} \; \equiv \; \mathbb{V}_t\left[r_s\right] \; = \; \frac{\sigma_r^2}{\alpha}\,(1 - e^{-2\beta(s-t)}) \qquad \text{for } s \geq t.$$

Let $\{\tilde{r}_t\}$ denote the discrete valued process that approximates the continuous valued process $\{r_t\}$. Let $\overline{r_1} < \overline{r_2} < \overline{r_3}, \dots, \overline{r_M}$ denote the values that $\tilde{r}_t$ may take on. A method of selecting the values $\overline{r_1}$ and $\overline{r_M}$ is to let the absolute value of the difference between $\overline{r_1}$ $(\overline{r_M})$ and $r_{s|t}^{\mathrm{mean}}$ be a multiple m of the conditional variance $r_{s|t}^{\mathrm{var}}$. Mathematically,

$$\overline{r_1} \; = \; r_{s|t}^{\mathrm{mean}} - m \times r_{s|t}^{\mathrm{var}}; \qquad \overline{r_M} \; = \; r_{s|t}^{\mathrm{mean}} + m \times r_{s|t}^{\mathrm{var}}.$$

Let the remaining $\overline{r_k}$s be equispaced in the interval $[\overline{r_1}, \overline{r_M}]$ and denote $\triangle \overline{r} = \overline{r_j} - \overline{r_{j-1}}$ where $j \in \{2, 3, \dots, M\}$.

We set one period in the Markov-Chain to be T years. The probability of making a transition from node $\overline{r_j}$ to node $\overline{r_k}$ in T years is calculated as follows. For each node j and for all

$n \in \{0, 1, \ldots\}$

$$
\begin{aligned}
\pi_{jk} &= \operatorname{Pr}[\tilde{r}_{(n+1)T} = \overline{r_k} \,|\, \tilde{r}_{nT} = \overline{r_j}] \\[2mm]
&= N\left[\frac{\overline{r_k} - r_{T|0}^{\text{mean}} + \triangle\bar{r}/2}{\sqrt{r_{T|0}^{\text{var}}}}\right] - N\left[\frac{\overline{r_k} - r_{T|0}^{\text{mean}} - \triangle\bar{r}/2}{\sqrt{r_{T|0}^{\text{var}}}}\right] \qquad \text{if } k \in \{2, 3, \ldots, M-1\} \\[2mm]
&= N\left[\frac{\overline{r_k} - r_{T|0}^{\text{mean}} + \triangle\bar{r}/2}{\sqrt{r_{T|0}^{\text{var}}}}\right] \qquad \text{if } k = 1 \\[2mm]
&= 1 - N\left[\frac{\overline{r_k} - r_{T|0}^{\text{mean}} + \triangle\bar{r}/2}{\sqrt{r_{T|0}^{\text{var}}}}\right] \qquad \text{if } k = M.
\end{aligned}
$$

Intuitively, the approximation works for the following reason. As the number of nodes M increase, the conditional distribution of $\tilde{r}_{(n+1)T}|\tilde{r}_{nT} = \overline{r_j}$ will closely approximate that of $r_{(n+1)T}|r_{nT} = \overline{r_j}$ in the sense of weak convergence.

With the discrete approximation of r_t in place, the tax benefit, bankruptcy cost and transaction cost can be written as

$$
\begin{aligned}
\overline{\mathbf{tb}(\bar{\mathbf{r}}, \zeta, \mathbf{T}, \mathbf{V_0})} &= V_0 \times \left[\, \overline{ntb_1} \,, \overline{ntb_2} \,, \ldots, \overline{ntb_M} \,\right]' \qquad \text{with} \quad \overline{ntb_j} \doteq ntb(r_j, \zeta, T), \\
\overline{\mathbf{bc}(\bar{\mathbf{r}}, \zeta, \mathbf{T}, \mathbf{V_0})} &= V_0 \times \left[\, \overline{nbc_1} \,, \overline{nbc_2} \,, \ldots, \overline{nbc_M} \,\right]' \qquad \text{with} \quad \overline{nbc_j} \doteq nbc(r_j, \zeta, T), \\
\overline{\mathbf{tc}(\zeta, \mathbf{V_0})} &= V_0 \times \left[\, \overline{ntc_1} \,, \overline{ntc_2} \,, \ldots, \overline{ntc_M} \,\right]' \qquad \text{with} \quad \overline{ntc_j} \doteq ntc(\zeta).
\end{aligned}
$$

Note that $\overline{\mathbf{tb}(\bar{\mathbf{r}}, \zeta, \mathbf{T}, \mathbf{V_0})}, \overline{\mathbf{bc}(\bar{\mathbf{r}}, \zeta, \mathbf{T}, \mathbf{V_0})}, \overline{\mathbf{tc}(\zeta, \mathbf{V_0})}$ is a $M \times 1$ vector.

Scalability

Consider for the moment that horizon is finite so that the firm can only issue debt for $N-1$ periods. Economically, the firm exogenously dies at time NT. Then, the present value of tax benefits at time $(N-1)T$ are

$$
\overline{TB_{j;N-1}} \doteq \overline{TB_{N-1}(r_j, \zeta, T, V_{(N-1)T})} = V_{(N-1)T} \times ntb_j.
$$

The following derivation shows that the tax benefit at time $(N-2)T$ is linear in $V_{(N-2)T}$. Define τ_{N-2} as the first passage time when the unlevered firm value breaches the default threshold after firm issues debt at time $(N-2)T$. Using backward induction, the present

value of tax benefits at time $(N-2)T$ are

$$
\begin{aligned}
\overline{TB_{j;N-2}} &\doteq \overline{TB_{N-2}(r_j,\zeta,T,V_{(N-2)T})} \\
&= V_{(N-2)T} \times ntb_j \\
&+ \sum_{k=1}^{M} \pi_{jk}\mathbb{E}_{N-2}\left[\overline{TB_{j;N-1}}\,\mathbb{I}_{\tau_{N-2}>T}\,e^{\int_{(N-2)T}^{(N-1)T} -r_u\,du}|r_{(N-1)T}=r_k\right]
\end{aligned}
\qquad (2.11)
$$

The first term of equation (2.11) is the present value of tax benefit from the debt issued at $(N-2)T$. The second term of equation (2.11) is the present value of tax benefit from the subsequent debt issuance at $(N-1)T$ conditional upon surviving until $(N-1)T$. In subsection 5.7 of the Appendix, we show that

$$
\overline{TB_{j;N-2}} = V_{N-2}\,e^{-yT}\,ntb_k\,H_{jk}
\qquad (2.12)
$$

where

$$
H_{jk} \doteq H(r_j,r_k,T) = \mathbb{E}_{N-2}\left[e^{\int_{(N-2)T}^{(N-1)T} -0.5\sigma_v^2 du + \int_{(N-2)T}^{(N-1)T} \sigma_v dW_{vu} du}\,\mathbb{I}_{\tau_{N-2}>T}|r_{(N-1)T}=r_k\right].
$$

We solve for H_{jk} by numerical simulation. In the appendix, we also give a closed form expression that provides a very good approximation of H_{jk}. Equation (2.12) says that the present value of tax benefit at time $(N-2)T$ is a linear function of the unlevered firm value $V_{(N-2)T}$. By induction, the present value of tax benefit at time zero is linear in V_0. Mathematically, it means that the tax benefits at time zero can be written as

$$
\overline{TB_{j;0}} = V_0 \times ftb_j
$$

where ftb_j is some function independent of the firm value V_0. The function ftb_j is the normalized present value of tax benefits per unit of unlevered firm value. By letting $N \to \infty$, we apply a fixed point argument to $\overline{TB_{j;0}}$, yielding

$$
\overline{TB_{j;0}} = \overline{tb_j} + \sum_{k=1}^{M} \pi_{jk}\,H_{jk}\,\overline{TB_{k;0}}.
$$

In matrix notation, we have that

$$
\begin{pmatrix} \overline{TB_{1;0}} \\ \overline{TB_{2;0}} \\ \vdots \\ \overline{TB_{M;0}} \end{pmatrix}
=
\begin{pmatrix} \overline{tb_1} \\ \overline{tb_2} \\ \vdots \\ \overline{tb_M} \end{pmatrix}
+
\begin{pmatrix}
\psi_{11} & \psi_{12} & \cdots & \psi_{1M} \\
\psi_{21} & \ddots & \cdots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{M1} & \cdots & \cdots & \psi_{MM}
\end{pmatrix}
\begin{pmatrix} \overline{TB_{1;0}} \\ \overline{TB_{2;0}} \\ \vdots \\ \overline{TB_{M;0}} \end{pmatrix}
$$

where $\psi_{jk} = \pi_{jk}H_{jk}$. Define $\mathcal{I}$ as the identity matrix with M dimensions, and Ψ as a matrix with elements ψ_{jk}, then we have that

$$
ftb_j = \frac{ntb_j}{\mathcal{I} - e^{-yt}\Psi}.
$$

Analogously, the normalized transaction costs and bankruptcy costs are

$$ftc_j \; = \; \frac{ntc_j}{\mathcal{I} - e^{-yt}\Psi}; \qquad fbc_j \; = \; \frac{nbc_j}{\mathcal{I} - e^{-yt}\Psi}.$$

Finally, the total levered firm value equals the after tax unlevered firm value, plus the value of tax benefit, less the value of bankruptcy and transaction cost. The firm value for a given $\bar{r}_j$ is given by

$$\overline{TV_{j;0}} \; = \; V_0 \times [(1 - \theta) + ftb_j \; - \; fbc_j \; - \; ftc_j]. \tag{2.13}$$

The firm chooses maturity T to maximize the total firm value in equation (2.13).

2.6 Quantitative analysis

Before we analyze the optimal debt maturity, it is useful to analyze the levered firm value, tax benefit, bankruptcy cost and transaction cost for an arbitrary maturity. In Figures 2.7, 2.8 and 2.9, we plot the normalized value of each component as a function of maturity for positive, negative and zero values of the term spread, respectively. For example, bankruptcy cost of \$0.1 means that the bankruptcy cost is 10 cents per dollar of unlevered firm value. These graphs show the tradeoff faced by the firm while choosing optimal debt maturity. We analyze each component sequentially.

First, consider transaction costs, as shown in the dark solid line with square markers in the three figures. As expected, the transaction costs decrease with maturity. This result is mechanical: as debt maturity increases, the need to rollover debt decreases, which in turn reduces transaction costs. The dependence of the transaction costs on the term spread is more interesting. From the graphs, it is evident that the transaction costs are almost independent of the term spread as indicated by the small magnitude of the slope of these lines. From equation (2.7), the one period transaction cost is not a function of the short rate. However, from (2.13), it is clear that the transaction cost depends on the sample path of short rate $\{r_t\}$. It turns out that for choice of parameters within the range of economic interest, the dependence between the transaction costs and the short rate is insignificant.

Second, consider bankruptcy costs, as shown in the dotted solid line in the three figures. For each level of the term spread, bankruptcy costs approach zero as time to maturity decreases to zero. This result is expected, since in our model interest rates and firm value are driven by continuous Wiener processes. Therefore, the probability of bankruptcy smoothly reaches zero as the time to maturity decreases to zero (Duffie and Lando (2001)). In addition, observe that for a given maturity, bankruptcy cost is lowest when the term spread is negative and it is highest when the term spread is positive. The intuition for this result also follows directly from Figure 2.3. When the term spread is positive, that is when the short rate is low, the chance of breaching the default threshold is high.

Third, consider tax benefits, as shown in the dark solid line with round markers in the three figures. Two observations are in order. First, note that the slope of tax benefits is almost zero in all figures. Second, note that the level of tax benefits is almost the same for different term spreads. In addition to the slope, note that the level of the tax benefits are the same across different term spreads. From the figures, it is clear that tax benefits will play a minor role when firms optimize debt maturity. This evidence is consistent with Graham and Harvey (2001) whose survey results point out that CFO do not consider tax benefit while choosing debt maturity.
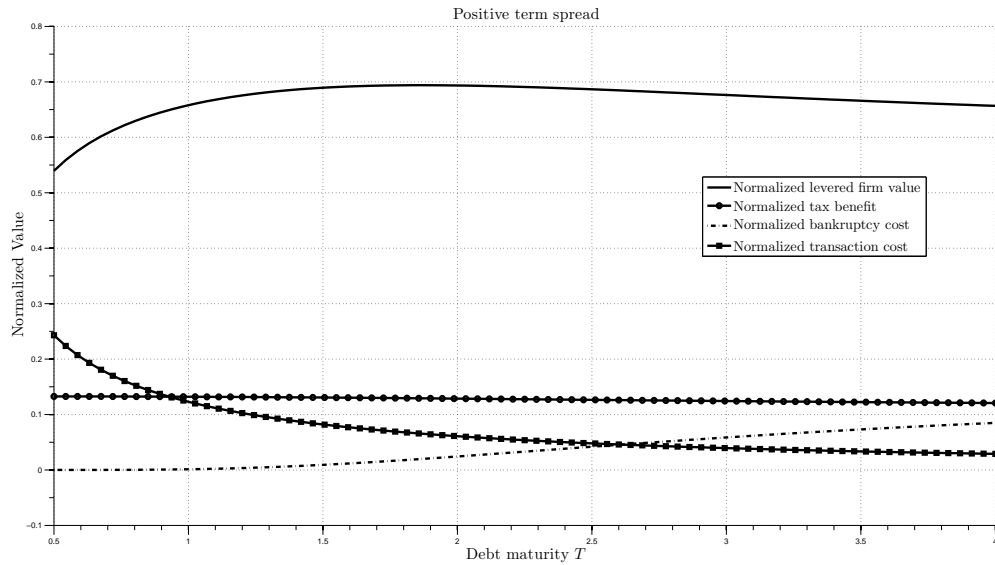
Figure 2.7: This figure shows the relationship between debt maturity and (i) the total firm value (ii) tax benefit (iii) bankruptcy cost, and (iv) transaction cost. The parameters are as follows: $\beta = 0.261$, $V_0 = 100$, $\alpha = 0.0716$, $\gamma = 0.5$, $\theta = 0.35$, $y = 0.05$, $P_0 = 40$, $\sigma_v = 20\%$, $\phi = 2\%$, and $r_0 = 2.16\%$. Note that the term spread is significantly positive since the short rate r_0 is significantly lower than the long rate α.

Lastly, with all the components in place, we analyze the effect of different parameters on the firm value next.

Effect of term spread on firm value

Consider firm values, as shown in the dark solid line in Figures 2.7, 2.8 and 2.9. The firm value is highest when the term spread is negative and it is lowest when the term spread is positive. This is also consistent with empirical evidence. Firm values are high prior to the beginning of a recession when the term spread is negative. Conversely, firm values are the lowest at the end of a recession when the term spread is positive. This result matches the literature concerning the predictability of equity returns as surveyed in the presidential address by John Cochrane (Cochrane (2011)).
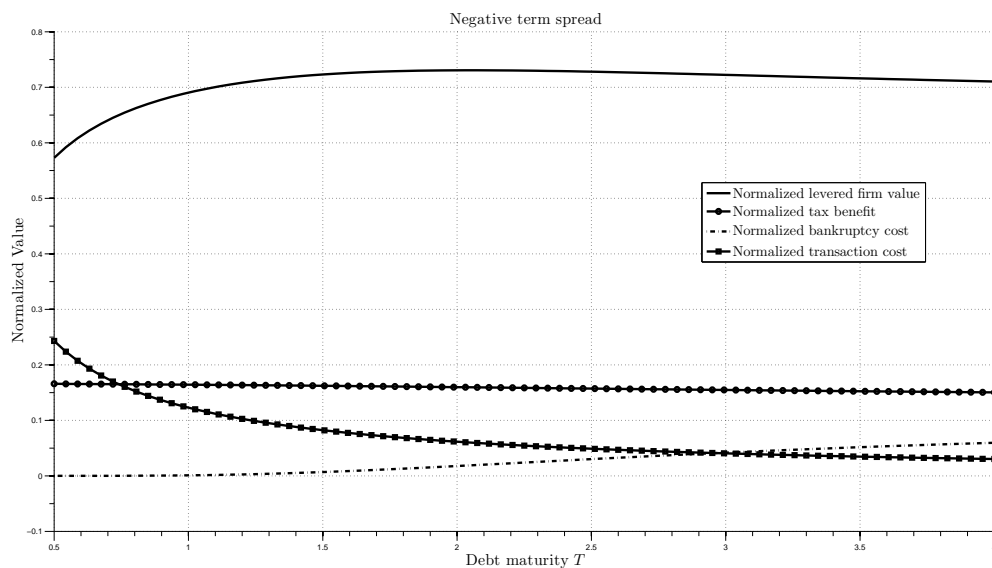
Figure 2.8: This figure shows the relationship between debt maturity and (i) the total firm value (ii) tax benefit (iii) bankruptcy cost, and (iv) transaction cost. The parameters are as follows: $\beta = 0.261$, $V_0 = 100$, $\alpha = 0.0716$, $\gamma = 0.5$, $\theta = 0.35$, $y = 0.05$, $P_0 = 40$, $\sigma_v = 20\%$, $\phi = 2\%$, and $r_0 = 12.16\%$. Note that the term spread is significantly negative since the short rate r_0 is significantly greater than the long rate α.
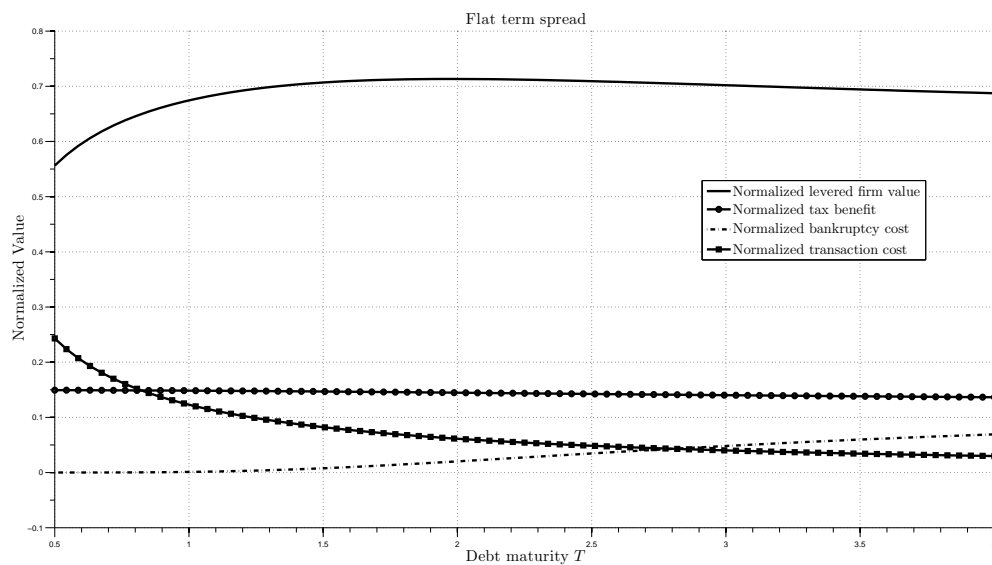
Figure 2.9: This figure shows the relationship between debt maturity and (i) the total firm value (ii) tax benefit (iii) bankruptcy cost, and (iv) transaction cost. The parameters are as follows: $\beta = 0.261$, $V_0 = 100$, $\alpha = 0.0716$, $\gamma = 0.5$, $\theta = 0.35$, $y = 0.05$, $P_0 = 40$, $\sigma_v = 20\%$, $\phi = 2\%$, and $r_0 = 7.16\%$. Note that the term spread is approximately zero since the short rate r_0 is equal to the long rate α.

Effect of term spread on optimal debt maturity

Figure 2.10 plots the optimal maturity as a function of the term spread for different values of firm leverage. The solid line plots the optimal maturity for highly levered firms; the solid line with round markers shows the optimal maturity for medium levered firms; and the dotted dashed line shows the optimal maturity for low levered firms. From the slightly decreasing nature of the plots, it is clear that optimal maturity is a decreasing function of the term spread. This result is consistent with the empirical findings of Barclay and Smith (1995), Guedes and Opler (1996), Ozkan (2000), Graham and Harvey (2001), Faulkender and Petersen (2006), Julio, Kim, and Weisbach (2008), and Chen, Xu, and Yang (2012).
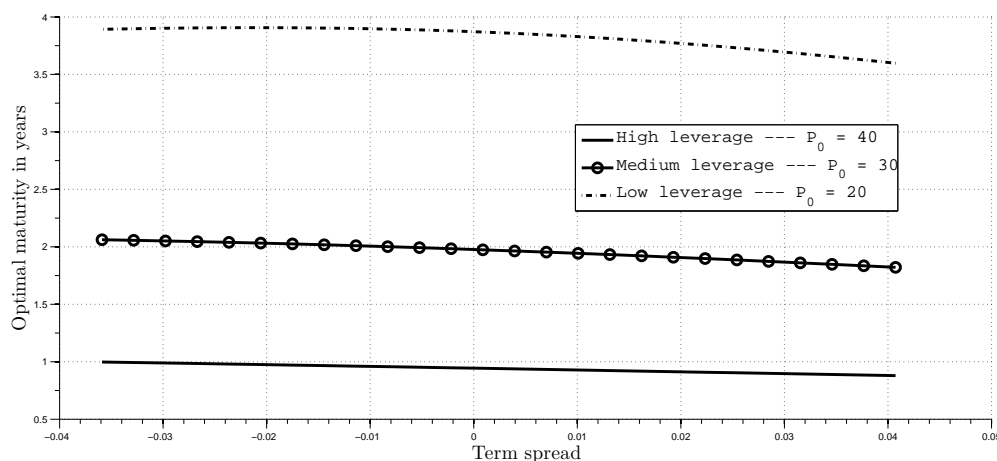
Figure 2.10: This figure shows optimal maturity as a function of the term spread for different firm leverages. The parameters are as follows: $\beta = 0.261$, $V_0 = 100$, $\alpha = 0.0716$, $\gamma = 0.5$, $\theta = 0.35$, $y = 0.05$, $\sigma_v = 20\%$, $\phi = 2\%$, and $r_0 = 7.16\%$.

Effect of leverage on optimal debt maturity

From Figure 2.10, it is also clear that optimal debt maturity is inversely related to leverage. For example, for a given term spread, optimal maturity of highly levered firms is lower than optimal maturity of low levered firms. That is, highly levered firms choose lower debt maturity than low levered firms. This is also consistent with the evidence in Barclay and Smith (1995), and Julio, Kim, and Weisbach (2008).

Effect of transaction costs, firm volatility and correlation on optimal debt maturity

Table 2.3 summarizes the effect of a small positive change in either transaction cost parameter ϕ, firm volatility σ_v or correlation ρ on the various components of the firm value.

Not surprisingly, an increase in the transaction cost parameter ϕ lowers the firm value and increases optimal debt maturity. Less obvious is the comparative statics with firm volatility. An increase in firm volatility σ_v increases the probability of bankruptcy for a given leverage, which increases bankruptcy costs. An increase in the probability of bankruptcy also lowers the chance of future debt issuances, which in turn lowers both tax benefits and transaction costs. However, the increase in bankruptcy costs outweighs the decrease in transaction costs. Therefore, firms optimally choose to decrease maturity. The same intuition holds for the comparative statics for correlation ρ.

Change in variable	Tax benefit	Bankruptcy cost	Transaction cost	Firm value	Optimal maturity
Transaction cost ϕ	0	0	+	-	+
Firm volatility σ_v	-	+	-	-	-
Correlation ρ	-	+	-	-	-

Table 2.3: This table shows the comparative statics.

Discussion of the speed of adjustment toward the target leverage ratio

A primary assumption of our model is that firms adjust their capital structure toward a target leverage ratio. For a wide variety of parameters, we show that firms adjust their capital structure between 1-3 years. Specifically, our model results indicate that highly levered firms rollover their debt every year, while low levered firms rollover their debt every 2-3 years. Regardless the leverage ratio, firms are active in adjusting their capital structure. This result is in contrast with Strebulaev (2007) who shows that firms are inactive in adjusting their capital structure most of the time.

The fact that in our results firms are active in adjusting their capital structure is consistent with Faulkender and Smith (2011) who show that overlevered firms close more than 70% of the gap between actual and target leverage ratio upon realizing positive cashflows. Even firms with near zero cash flow realization close the gap between actual and target leverage ratio by 25%. Flannery and Rangan (2006) show that a typical firm closes about one-third of the gap between actual and target leverage ratio each year. Marcus (1983), Jalilvand and Harris (1984), Leary (2002), Leary and Roberts (2005) and Welch (2012) also show that managers are active in adjusting their capital structure.

We conclude this section by highlighting the following observation. The present literature overwhelmingly attributes the fact that firms issue short term debt during bad times towards

the validation of either agency theory, such as in Myers (1977), asymmetric information, such as in Diamond (1991) and Flannery and Rangan (2006), or liquidity risk, such as in Chen, Xu, and Yang (2012). Our results show that this empirical evidence is also consistent with the trade-off theory.

2.7 Conclusion

In this paper, we explain the link between debt maturity and term spread using the tradeoff theory of capital structure. Specifically, we show that:

1. Firms issue shorter term debt when term spread is positive, and they increase maturity as term spread decreases.

2. Firms are incredibly active in adjusting their capital structure.

3. Highly levered firms issue shorter term debt as compared with low levered firms.

4. High volatility firms issue shorter term debt as compared with low volatility firms.

5. Firms choose maturity by balancing bankruptcy costs and debt rollover costs. Tax benefits play a minor role in debt maturity choice.

We develop a model of optimal maturity structure with a Vasicek (1977) interest rate process. Valuation formulas are obtained in semi-closed form. We use a novel fixed-point argument with stochastic interest rates to obtain the total tax benefit, bankruptcy cost, and transactions cost for a dynamic model with an infinite number of debt issuances.

We can extend the model in three major directions. First, we can endogenize the cashflows generated by the firm. This allows us to express Tobin's Q as a function of the term spread. Second, we can add one more factor for interest rates to decompose debt maturity as a function of both term spread and level of interest rates. Lastly, we can incorporate asset substitution and other agency related frictions by endogenizing the default boundary as in Leland (1994).

Chapter 3

Form Follows Finance

3.1 Introduction

The tradition in urban economics is to treat building height as a tradeoff between rent and accessibility. In spatial equilibrium where no individual has an incentive to move, Alonso (1964), Mills (1967), and Muth (1964) show that commuting cost differences cause differences in the price of living space. This price compensating differential leads to two implications regarding building height. First, the buildings are tall near the centers of most cities. Second, building heights positively correlate with population. That is, more populated cities have taller buildings than other less populated cities. The price, quantity and height implications resulting from this simple tradeoff forms the basis of the so called "standard model" in urban economics. In a handbook chapter, Brueckner (1987) gives an excellent summary of the implications of the standard model.

Upon first glance, a casual introspection leads to a validation of the standard model. Helsley and Strange (2008) show the correlation between population and height of the tallest building is 0.65 for the 20 largest U.S. cities. However, the standard model ignores one key factor: debt financing from the capital markets. For example, the words "debt", "equity" , "capital markets" , and "uncertainty" are absent from a word search in the handbook chapter of Brueckner (1987).

In this paper, we analyze the effect of debt on the building heights. Capital markets are important for two reasons. First, casual observation leads to the conclusion that almost all commercial buildings are debt financed. Feng, Ghosh, and Sirmans (2007) show that the average debt to asset ratios of REITS is around 65% in spite of the fact REITS pay out 90% of their earnings. Second, the standard model is static in nature, while the real estate markets are prone to boom and bust cycles. Therefore, it is not too difficult to imagine that building heights are dependent on the opportunity cost of debt.

Figure 3.1 shows the relationship between the number of completed buildings in New York and the difference between AAA and BBB spreads of nonfinancial corporate bonds. We use the difference between AAA and BBB spreads of corporate bonds as a measure of the opportunity cost of debt. The reason is as follows. The difference between AAA and BBB spreads measures the magnitude of default premium charged by the market after controlling for liquidity and tax factors. For example, consider the year 2007 prior to the onset of the past recession. The difference between AAA and BBB spread is low which indicates that debtors do not pay much more to increase leverage. That is, the marginal cost of increasing debt is low. The negative relationship between the number of completed buildings and opportunity cost of debt is clear. During times when the opportunity cost of debt is low, there is an increase in the number of completed buildings.
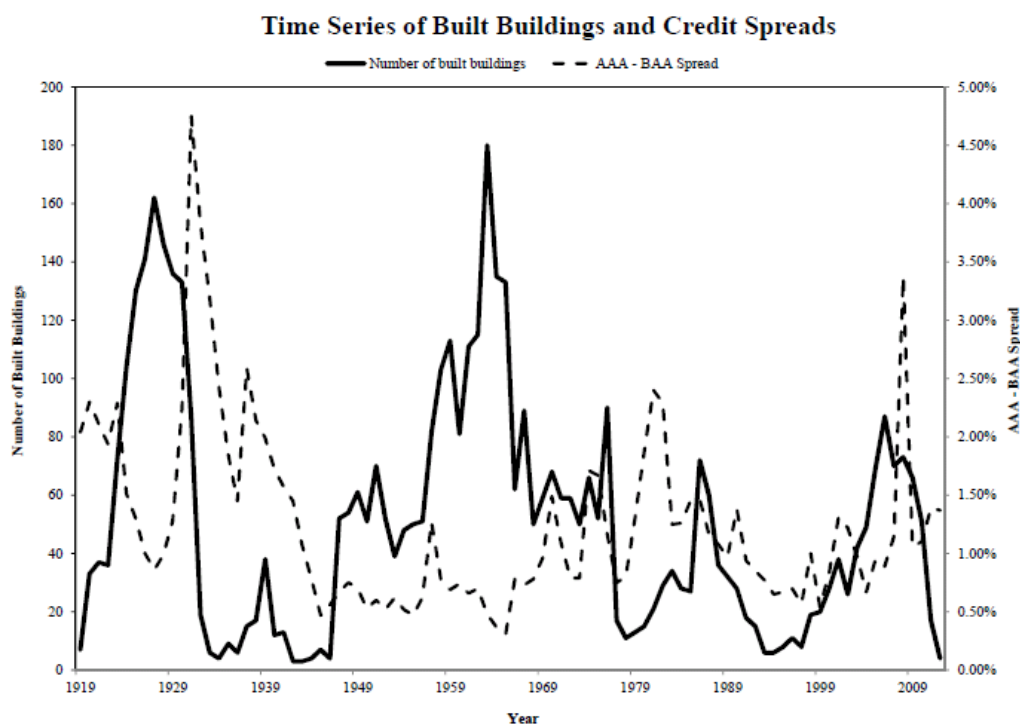
Figure 3.1: This figure shows the time series of the number of completed buildings in New York (solid line) from 1919 to 2012. The data is from Skyscrapper city.com. The dashed line is the difference between AAA and BBB nonfinancial corporate bonds.

Figure 3.2 shows the relationship between the tallest constructed building in New York and AAA and BBB spreads. Again, the negative relationship is clear. During times when the opportunity cost of debt is low, developers build taller buildings.

The choice of debt financing affects the skyscraper owner's height decision in the following way. Consider the situation when the owner is a monopolist and she does not pursue debt financing. Furthermore, future rents are uncertain. An increase in skyscraper height

Time Series of Maximum Height and Credit Spreads

── Maximum Number of Floors　　− − AAA - BAA Spread
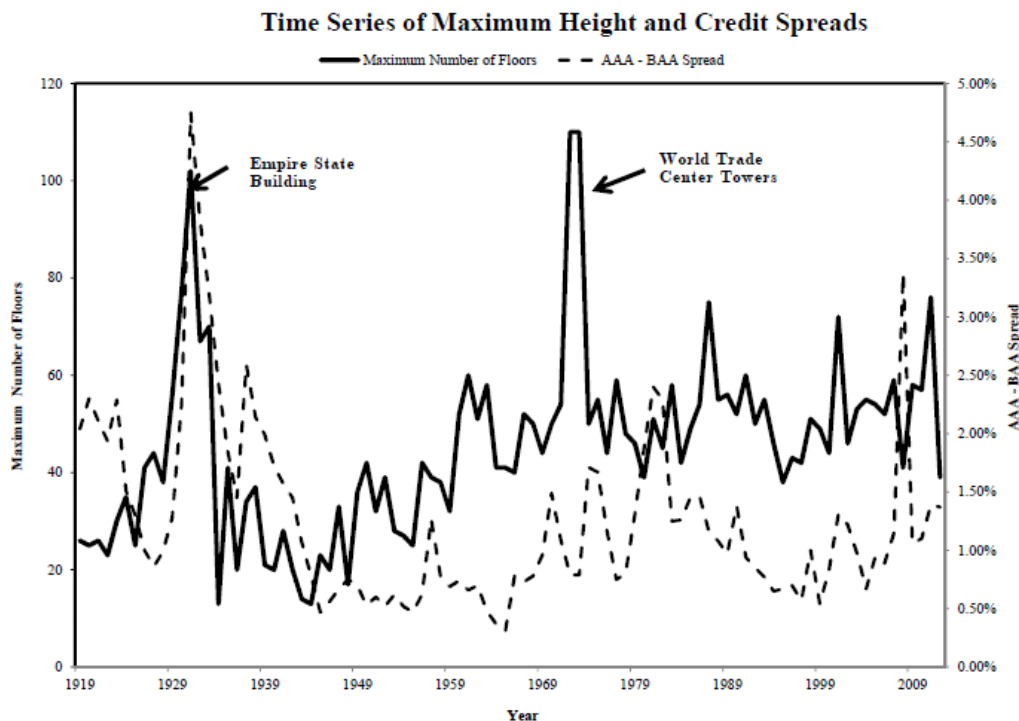
Figure 3.2: This figure shows the time series of the tallest constructed building in New York (solid line) from 1919 to 2012. The data are from Skyscrapper city.com. The dashed line is the difference between AAA and BBB nonfinancial corporate bonds. .

increases the supply of rental space, which in turn decreases expected rental prices. An increase in height also increases construction costs. Therefore, the owner chooses height to equate expected marginal revenues and expected marginal costs. Now, suppose the owner pursues debt financing. She chooses the amount of debt and height to maximize her equity value. Optimally, she chooses the amount of debt and height that raise returns in the good states. In the bad states, the owner has the option to declare bankruptcy. Therefore, the owner chooses height to equate expected marginal revenues and expected marginal costs only in the good (solvent) states. Debt financing leads the owner to be more aggressive — the owner builds taller than before.

Crucial to the setup of our model is the idea that an increase in height increases the supply of rental space, which in turn decreases expected rental prices. That is, height takes the role of quantity in a standard Cournot model. With this interpretation, we use the framework of Brander and Lewis (1986) to analyze financing and height decision faced by the owner. Other work that focuses on the link between financing decisions and product market include Brander and Lewis (1988), Maksimovic (1988), Bolton and Scharfstein (1990), Rotenberg and Scharfstein (1990), Chevalier (1995), Campello (2006), and Clayton (2009).

Our contributions are twofold. Our primary contribution is to show the importance of debt financing with a case study. We analyze a letter written by John Raskob, who was the original owner of the Empire State Building (hereafter referred to as ESB), to his banker Louis Kauffman. The letter, dated August 1929, is a pro forma analysis of ESB. Prior to construction, Raskob analyzes two projects: a 55-storey ESB and an 80-storey ESB. With the use of higher leverage, Raskob convinces Kauffman to proceed with an 80-storey project in part due to its higher return on capital. The letter highlights two salient features that form the basis of our model. First, the letter explicitly highlights how the limited liability feature of debt affects the decision height. Second, the letter implicitly highlights sequential nature of the decision making process of Raskob. Specifically, Raskob chooses how much to borrow in the first stage. Afterward, in the second stage, he chooses how tall to build. The sequential decision making process is present in any model that links product market competition and debt financing. With this historic letter, we are the first to validate the sequential nature of the decision making process explicitly.

Our second contribution is to provide micro foundations of the economics of a tall building with debt financing. The standard model assumes that real estate markets are perfectly competitive, whereas we relax that assumption. In the textbook model of Cournot competition, firms produce more quantity than what is optimal in autarky. Similarly, imperfect competition leads skyscraper owners to build taller than what is optimal in autarky. Helsley and Strange (2008) use behavioral preferences like ego along with imperfect competition to show that competition leads owners to construct a taller building.

Figure 3.3 shows the effect of both debt finance and imperfect competition on height. As explained before, both debt financing and imperfect competition separately cause the skyscraper owner to build taller. The shaded gray region shows the change in height from either increasing debt or increasing competition. The consequences of the last column are most surprising. Debt and imperfect competition taken together magnify the aggressive stance of the skyscraper owner. As a result, the skyscraper owners build much taller than what is optimal in autarky with 100% equity financing.

The paper is organized as follows. In the next section, we analyze the pro-forma analysis of Raskob while deciding the height of the ESB. In Sections III, IV and V, we analyze the model and provide some comparative statics. In Section VI, we provide concluding remarks.

3.2   Case study – Empire State building (ESB)

This section describes the decision making process behind the construction of the ESB. Specifically, we analyze a letter [see Figure 3.4] written by the original ESB owner, John Raskob, to his banker Louis Kauffman. This letter (dated August 28, 1929) is the primary motivation of the paper and forms the basis of the theoretical model. The letter explicitly
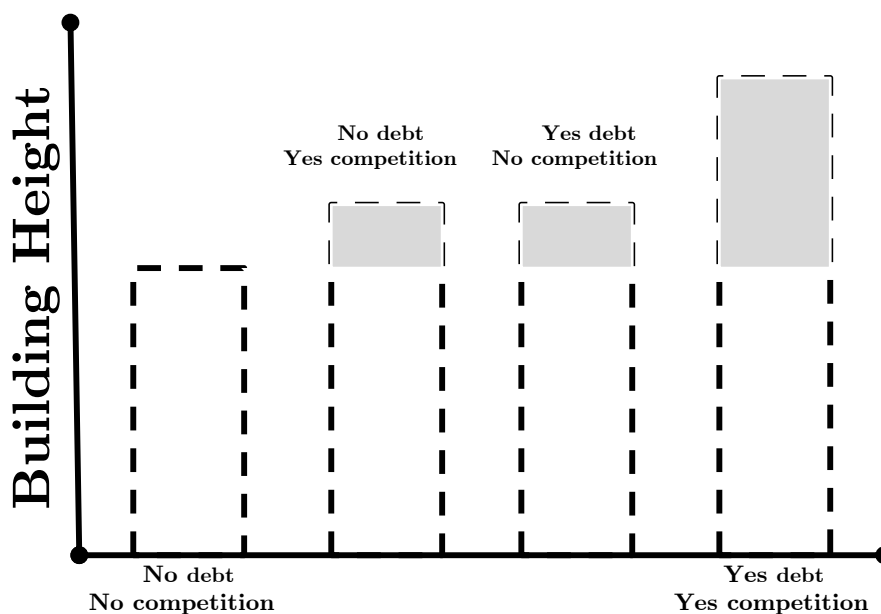
Figure 3.3: This figure shows the effect of debt and competition on skyscraper height.

points out the role of debt as well as other germane features associated with skyscraper development.

Background of the Letter

In 1928, architect and developer, Floyd Brown paid 14 million dollars for the Waldorf-Astoria Hotel, the site of the future ESB. The project was debt financed. His plan was to demolish the hotel and construct a 50-story, two million square feet, mixed-use building. Brown invested one hundred thousand dollars of his own money and borrowed $900.00 K from Chatham Phoenix National Trust Company; the sum of one million dollars went towards the first mortgage payment. Subsequently, Brown published his design in the 1928 December issue of Real Estate Record and Builders' Guide. This was a standard practice at that time. The publication served as an advertisement to attract potential investors. Unfortunately, investors did not flock to his proposal and consequently he failed to make his next mortgage payment of 1&1/2 million dollars. Chatham National took over the project. Louis G. Kauffman — president of Chatham National — approached his friend, John Raskob, with a potential opportunity (around 1928- 1929).

After some due diligence, Raskob sends a letter to Kauffman that outlines his economic

motivation behind the ESB height preference.

Features of the letter

The letter presents a balance sheet and a cash flow statement for two potential projects: a 55 storey building and a 80-storey building. We focus on four features: the capital structure of ESB, construction costs, operating profits and commitment role of debt. We analyze each feature next.

Capital Structure underlying ESB

Upon first glance, the capital structure is quite intricate. The balance sheet involves the use of senior and subordinated debt, preferred and common stock. First, consider the capital structure for the 55-storey building. The total construction cost of $45.00 MM is financed by two mortgages that account for $35.00 MM and preferred stock that accounts for the rest. The resulting debt to asset ratio is 77.77%. Now consider the income statement. After netting the revenues, operating costs and the interest payments, the cash flow accruing to common equity is actually negative (-$465.00 K).

Second, consider the capital structure for the 80-storey building. The addition of extra 25 stories requires an additional $5.00 MM, which is entirely financed by the two mortgages. As a result, debt to asset ratio increases to 80.00%. Now, consider the income statement. The net cash flow accruing to common equity is zero because of the higher revenues from a taller building.

The Return on Capital (ROC) for the 55-story building is 11.38% ($5.12MM / $45.00MM) while the ROC for the 80-storey building is 12.60% ($6:30MM / $50.00 MM). Higher ROC for the 80-story building is due to its higher leverage. Higher return on assets probably convinced both Raskob and Kauffmann to proceed with the 80-story building.

Note that the opportunity cost of debt is also low. An increase in the debt to asset ratio for the 80-storey building does not lead to an increase in the required yield. To summarize, Raskob, the equity holder, implicitly understands the limited liability nature associated with the debt financing. He chooses to be more aggressive and build tall because of access to cheap credit.

Commitment role of debt

Note the sequential two-stage procedure behind the choice of debt financing and height. The financial decision of how much to borrow is made in the first stage and the height decision is made in the second stage. The first stage is implicit in the letter while the second stage involving height choice is explicit. The two-stage decision-making process imparts debt an

	55 Storeys	80 Storeys
Land - - - - - - - - - - - - - -	-$16,000,000	Land - - - - - - - - -$16,000,000
29,000,000 on ft. @ $1.00 - - -- - - -	29,000,000	34,000,000 cu. ft. @ $1.00 (25 addtl storeys 80 x 240 — 34,000,000
Total cost - - - - -	45,000,000	50,000,000
1st Mtg (5½% - 2% S.F.) - - - - - -	25,000,000	27,500,000
Balance - - - - - - -	20,000,000	22,500,000
2nd Mtg (6½% with 20% of Com.) (Stk as bonus)	10,000,000	12,500,000
Balance - - - - - - -	10,000,000	10,000,000
Pfd. Srk (7% with 80% of Com.bonus)	10,000,000	10,000,000

I N C O M E

	55 Storeys	80 Storeys
1,750,000 sq. ft. @ $3.25 - - - - -	5,690,000	1,750,000 sq. ft. @ $3.25 — 5,690,000; 330,000 " " @ $4.00 — 1,310,000; 7,000,000
Vacancies 10 % - - - - - - - - -	570,000 / 5,120,000	700,000 / 6,300,000
Op. Exp. 1,750,000 sq. ft. @ 75¢	1,312,500	2,080,000 Sq.Ft.@ 75¢ — 1,560,000
Taxes 40,000,000 " @ $2.66	1,047,500	44,000,000 " @$2.66 — 1,170,000
Total Expense - - -	2,360,000	2,730,000
Bal. for Capital - - - - - - - - -	2,760,000	3,570,000
1st Mortg. Interest - - - - - - - -	1,375,000	1,510,000
	1,385,000	2,060,000
1st Mtg. S. F. 2% - - - - - - - -	500,000	550,000
	885,000	1,510,000
2nd Mtg. Int - - - - - - - - - -	650,000	810,000
Bal. for owners	235,000	700,000
$10,000,000 - 7% Pfd. Stk	700,000	700,000
Bal. for Com. Stk	465,000	0

Figure 3.4: This is a letter sent by John Raskob to Louis Kauffman. The letter describes the economic analysis for two cases: a 55 story building and 80 story building. The analysis takes into account debt costs, construction costs, operating costs, revenues, and vacancies. The letter is from Willis (1995).

important role — debt serves as a commitment device. In the jargon of game theory, by accessing the capital markets, Raskob publicly announces that he is going to debt finance ESB and thus he credibly commits to construct a taller building. The commitment role of debt is even more important in the presence of competition as shown below.

Competition

Competition is notably absent from the letter, but it cannot be ignored. This period in New York is commonly characterized by the phenomenon of "Skyscraper Race". Along with a historical account of the race, Helsley and Strange (2008) provide a game-theoretic analysis that considers the effect of competition on the choice of height. Particularly, in their setup, the owner of the tallest building receives some non-monetary benefit from owning the tallest building. The owner's ego is an example of such non-monetary benefit. As a result, owners build taller than what is optimal in autarky. Our model is similar in spirit but we abstain from non-monetary benefits. The owners in our setup are profit maximizers who use capital markets to maximize their equity. Due to imperfect competition, the skyscraper owner uses debt strategically to enhance its position in the product market.

Construction costs

Construction costs can be divided into two categories: fixed cost of $16.00 MM that predominantly involves land acquisition and variable building cost of $1.00 per sq. ft. that includes the cost of demolition, architect and builder's commission, and material (steel) costs.

First consider the fixed cost of acquiring land. Raskob assumes a cost of $191 per square foot for a 197 ft × 425 ft plot of land that is a few blocks south of the Grand Central Station. This is in line with the estimate by Clark and Kingston (1930) [hereby referred to as CK] who estimate a cost of $200 per square foot of similar size for a corner lot near Grand Central Station.

Next consider the variable building cost. Raskob assumption of a constant average cost of $1.00 per sq.ft. seems to be simplistic (and generous) in comparison to CK's analysis. Figure 3.5 presents the different components of the costs. For example, the table shows that to provide one square foot of rentable space, the structural-steel cost for a 50-story building is $1.15, whereas the same cost for a 75-story building is $1.72. The additional cost is due to the need to reinforce the weight incurred by additional stories. Mathematically, it is clear that construction costs are convex. In our model, we assume that construction costs are indeed convex.[1] Figure 3.6 provides additional details of the construction costs as a function of building height (during 1930s).

---

[1]Ironically, it turned out that ESB was constructed under budget. Due to the great depression, the price of steel plummeted. In our model, we ignore variable building cost uncertainty.

Component Factors	8 Stories	15 Stories	22 Stories	30 Stories	37 Stories	50 Stories	63 Stories	75 Stories
Excavations and Foundations.	$0.738	$0.476	$0.399	$0.413	$0.435	$0.438	$0.449	$0.415
Structural Steel.	0.717	0.798	0.869	0.952	1.009	1 151	1.351	1 726
Concrete Floors.	0.608	0.580	0.575	0.572	0.583	0.603	0.625	0.642
Permanent Interior Partitions.	0.224	0.233	0.256	0.274	0.286	0.321	0.356	0.386
Brickwork	0.177	0.252	0.324	0.369	0.410	0.471	0.515	0.551
Exterior Finish.	0.397	0.497	0.515	0.531	0.517	0.508	0.506	0.515
Roofing.	0.060	0.057	0.048	0.040	0.037	0.036	0.032	0.030
Windows and Glazing.	0.191	0.183	0.195	0.201	0.209	0.221	0.230	0.239
Interior Finish and Trim	1.821	1 746	1 785	1 804	1.803	1 848	1.894	1 910
Mechanical Equipment	1 765	1.813	2.054	2.423	2.699	2.975	3.269	3.623
(a) Elevators and Elevator Fronts.	0.614	0.651	0.868	0.926	1 104	1.245	1.408	1.656
(b) Plumbing, Drainage and Water Supply	0.434	0.500	0.532	0.567	0.641	0.667	0.717	0.765
(c) Electric Light and Power, Wiring, etc.	0.269	0.305	0.323	0.352	0.364	0.384	0.402	0.419
(d) Heating and Ventilating.	0.448	0.356	0.331	0.578	0.590	0.679	0.742	0.785
Tenants' Changes.	0.594	0.662	0.68	0.700	0.701	0.699	0.693	0.688
Miscellaneous	0.397	0.286	0.244	0.309	0.280	0.256	0.238	0.227
Plant and General Conditions.	0.481	0.425	0.414	0.429	0.436	0.448	0.467	0.492
Extras and Contingencies.	0.230	0.224	0.223	0.218	0.217	0.218	0.220	0.219
Total Direct Labor and Material Cost	8.400	8.232	8.588	9.248	9.622	10.191	10.844	11.665
Total Building Cost.	9.289	9.098	9.463	10.100	10.514	11.089	11 728	12.589
Land Cost ($200 Per Sq. Ft.)	31.553	20.172	16.467	13.695	12.335	10.863	9.798	9.041
Total Carrying Charges (Per Sq. Ft.)	2.384	1 925	1 907	1 926	1 999	2.055	2.148	2.284
Total Investment (Per Sq. Ft.)	43.226	31 195	27 837	25 921	24.847	24.008	23.674	23 913
Total Cost Assignable to Land.	33 700	21 818	18.034	15.407	13 844	12 341	11 263	10.516
Total Cost Assignable to Building.	9.526	9.377	9.803	10.514	11.003	11.667	12.411	13.397

Figure 3.5: Detailed analysis of construction costs for different building heights. The table is from Clark and Kingston (1930).

BUILDING COST BY CHIEF COMPONENT FACTORS
(000 omitted)

Trades	8 Stories	15 Stories	22 Stories	30 Stories	37 Stories	50 Stories	63 Stories	75 Stories
Excavations and Foundations	$379	$382	$393	$481	$571	$653	$743	$744
Structural Steel	368	641	855	1,110	1,325	1,717	2,233	3,093
Concrete Floors and Finish	312	466	566	667	766	899	1,033	1,151
Permanent Interior and Partitions	115	187	252	320	376	478	589	692
Brickwork	91	202	319	430	539	703	851	988
Exterior Finish (Stone, Terracotta, Copper and Bronze, etc.)	204	399	507	619	679	757	836	922
Roofing	31	46	47	47	49	53	53	53
Windows and Glazing	98	147	192	234	274	329	381	429
Interior Finish and Trim	935	1,402	1,756	2,103	2,368	2,756	3,131	3,423
Mechanical Equipment	906	1,456	2,021	2,825	3,545	4,436	5,404	6,493
(a) Elevators and Elevator Fronts	315	523	854	1,080	1,450	1,856	2,328	2,967
(b) Plumbing, Drainage and Water Supply	223	402	523	661	·842	995	1,186	1,370
(c) Electric Light and Power, Wiring and Fixtures	138	245	318	410	478	572	664	750
(d) Heating and Ventilating	230	286	326	674	775	1,013	1,226	1,406
Tenants' Changes	305	532	676	816	920	1,042	1,146	1,232
Miscellaneous	204	230	240	360	368	382	394	407
Total	3,948	6,090	7,824	10,012	11,780	14,205	16,794	19,627
Plant and General Conditions	247	341	407	500	572	668	772	882
Extras and Contingencies	118	180	219	254	285	325	363	393
Total Direct Labor and Material Cost	4,313	6,611	8,450	10,766	12,637	15,198	17,929	20,902
General Contractor's Profit	259	390	465	538	632	684	770	836
Architects and Engineers	197	306	395	471	539	655	691	820
Total Building Cost	4.769	7,307	9,310	11,775	13,808	16,537	19,390	22,558

Figure 3.6: Detailed analysis of construction costs for different building heights. The table is from Clark and Kingston (1930).

Operating profits

Raskob calculates profits based on a rental rate assumption of \$3.25 - \$4.00 per sq.ft. and an operating cost of \$0.75 per sq.ft., which are in agreement with CK's analysis (see Figures 3.7 and 3.8). We deviate slightly from the setup and assume convex operating costs. Convexity is a technical assumption, which we use to ensure uniqueness of equilibrium.

66

GROSS RENTAL PER SQUARE FOOT OF NET RENTABLE AREA

Building	Average Sq. Ft. Rental for All Space	Average Sq. Ft. Rental for All Office Space	Average Sq. Ft. Rental for Ground Floor Shops
8 Stories.........	$3.54	$2.42	$13.24
15 Stories.........	3.46	2.77	13.80
22 Stories.........	3.54	2.96	14.48
30 Stories.........	3.59	3.11	14.95
37 Stories.........	3.62	3.22	15.45
50 Stories	3.75	3.36	16.98
63 Stories.........	3.82	3.48	17 60
75 Stories.........	3.86	3.54	18 23

Figure 3.7: Rental Rate assumption from Clark and Kingston (1930).

ANNUAL OPERATING EXPENSES PER SQUARE FOOT OF NET RENTABLE AREA

Building	Operating Expenses	Taxes	Depreciation	Total Expenses
8 Stories....	$0.61	$0.93	$0.19	$1.73
15 Stories....	0.60	0.67	0.18	1.45
22 Stories....	0.60	0.60	0.18	1.38
30 Stories....	0.62	0.56	0.20	1.38
37 Stories....	0.62	0.53	0.21	1.36
50 Stories....	0.63	0.52	0.22	1.37
63 Stories....	0.64	0.51	0.23	1.38
75 Stories....	0.68	0.52	0.25	1.45

Figure 3.8: Operating Cost assumption from Clark and Kingston (1930).

Vacancy Rates

Raskob also assumes vacancy rate of 10% which is also consistent with the historical vacancy rate in different neighborhoods on New York (see Figure 3.9). However, in retrospect, the vacancy rate is grossly underestimated. For example, in 1934 the vacancy rates increased to 25%—ESB was infamously called the "Empty State Building". Vacancy rates are also a result of the oversupply of rental space. For example, the total volume of office space constructed in New York between 1925 to 1929 was 17 MM sq.ft. and projects initiated by 1930 and completed in 1933 added another 13.00 MM sq.ft. (Willis (1995), page 170). We assume that demand for rental space is downward sloping and random. Therefore, states of nature with low demand imply that market-clearing prices are also low.

OFFICE BUILDING OCCUPANCY OF THE CENTRAL BUSINESS AREA IN NEW YORK CITY

Date of Survey	Financial Dist.	Insurance Dist.	City Hall Dist.	Grand Central Dist.	Plaza Dist.	Times Square Dist.	Columbus Circle Dist.	All Dist.
				Percent Occupied				
Feb. 1925	93.7	95.6	95.5	93.9				
May "	90.7	95.0	94.9	93.6				94.0
Oct. "	94.7	95.2	95.1	95.4				91.7
Jan. 1926	96.5	95.8	95.7	96.7	District not pre- viously surveyed	District not previously surveyed	District not pre- viously surveyed	95.0
Apr. "	96.9	95.3	95.1	96.8				95.2
Oct. "	96.9	95.4	95.0	96.8				95.5
Jan. 1927	97.2	95.5	95.4	95.1				95.8
Apr. "	97.5	95.4	95.2	95.2				95.8
Oct. "	96.7	95.5	94.7	87.1				95.7
Jan. 1928	96.8	95.5	94.8	88.9				92.0
May "	96.4	95.4	95.0	90.2				89.2
Oct. "	96.9	95.8	95.2	91.0	89.3		75.0	93.5
Jan. 1929	97.8	96.8	96.7	92.8	91.2		81.6	93.6
May "	98.8	97.8	94.5	91.5	92.0	District not previously surveyed	86.0	95.3
Oct. "	99.1	99.0	96.6	93.2	85.8		93.6	94.0
Jan. 1930	99.1	97.4	94.3	93.1	85.9		92.1	95.9
May "	95.6	97.0	88.8	88.0	75.3		91.8	95.7
Sept. "	95.5	97.7	90.5	88.1	71.7		92.8	91.1
Jan. 1931	96.1	97.4	90.7	88.5	68.7		90.6	91.2
May "	87.0	85.1	86.8	82.7	57.7		87.2	91.3
Sept. "	87.1	86.3	86.2	81.7	60.6		86.0	83.6
Jan. 1932	86.7	36.9	85.0	80.2	62.1		83.1	83.5
May "	83.1	83.3	83.1	78.1	55.0		80.7	82.7
Sept. "	82.1	82.6	82.8	77.1	56.5		80.5	79.7
Jan. 1933	81.4	82.9	82.2	76.9	59.3	73.3	77.1	79.4
May "	78.3	76.2	77.0	74.8	60.6	72.0	77.8	78.6
Sept. "	78.6	75.6	78.5	75.0	66.4	76.8	75.8	75.8
Jan. 1934	78.8	75.9	78.4	72.1	64.5	78.2	74.7	76.6
								75.2

Fig. 159 Office building occupancy of Central Business Area, New York, 1925 to 1934.

Figure 3.9: Office building occupancy of Central Business Area, New York, 1925 to 1934. The table is from Willis (1995), page 163.

Summary

CK conclude that a 63-story building provides the highest return (see Figure 3.10). Optimal height of ESB should be less than 63 stories as it is in a less desirable location than the Grand Central Station. Raskob's plan of 80 stories seems irrational in light of CK's analysis. However, CK's analysis ignores both debt financing and imperfect competition. In the model, which follows next, we show that the use of debt and imperfect competition indeed affected Raskob's decision to build taller than 63 stories. Furthermore, the decision to build taller is rational.
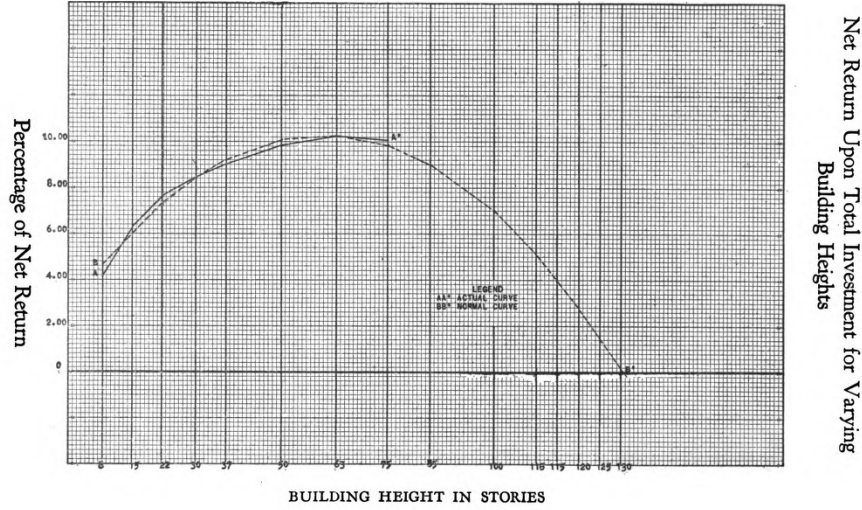
Figure 3.10: Comprehensive calculation of the rate of return of a hypothetical building situated near the Grand Central Station in NewYork. The table is from Clark and Kingston (1930).

3.3   Set up of the Model

Skyscraper financing

The decision-makers in our model are two competing agents, 1 and 2, who have an opportunity to build a skyscraper. The agents represent inside equity holders. Each agent i contributes some of her own financial capital, E^i, referred to as equity capital and borrows an amount B^i from the bond market to finance the skyscraper. The scale of the skyscraper is denoted q^i which represents the quantity of rentable space. For exposition, we normalize the units of q^i to one floor of a building. Therefore if $q^i > q^j$, then it means that agent i's skyscraper is taller than agent j's skyscraper.

The construction cost for a building with height q^i is

$$C^i(q^i; \gamma^i, \beta^i) = \gamma^i + c^i(q^i)\,\beta^i$$

where γ^i denotes the fixed costs, $c^i(q^i)$ denotes the variable building costs, and β^i controls the fraction of fixed costs relative to the variable building costs. Fixed costs may include the land acquisition costs as well as regulatory fees. Variable building costs may include labor costs and material (steel) costs that increase with the scale of the project. An increase in β^i increases the variable costs which may be due to an increase in the price of steel. Consistent with the Clark and Kingston's study, we assume that variable buildings costs are convex:

$$c_i^i \equiv \partial c^i / \partial q^i > 0; \qquad c_{ii}^i \equiv \partial^2 c^i / \partial q^{i2} > 0.$$

The financial restrictions imply that

$$C^i(q^i; \gamma^i, \beta^i) \leq E^i + B^i. \tag{3.1}$$

Equation 3.1 is due to financial constraints and it implicitly defines the height q^i as a function of fixed costs γ^i, variable costs β^i and financial investment $B^i + E^i$. Furthermore in equilibrium, equation 3.1 binds. From the differentiation rules of an inverse function, we have

$$dq^i/d\gamma^i \equiv q^i_{\gamma^i} < 0; \quad dq^i/d\beta^i \equiv q^i_{\beta^i} < 0; \quad dq^i/dB^i \equiv q^i_{B^i} = dq^i/dE^i \equiv q^i_{\gamma_i} > 0.$$

An increase in γ^i while holding financial investment $B^i + E^i$ and variable costs β^i constant reduces the amount left over to cover building costs $c^i(q^i)$ which in turn constricts height. Similar interpretations hold for other variables.

Operating cashflows from rental space

The demand for rentable space depends on random variable $z \in [\underline{z}, \overline{z}]$ with $\underline{z} > 0$, density function $f(z)$ and cumulative distribution function $F(z)$. Using $P(Q, z)$ to denote the downward sloping demand curve, agent i's revenues are

$$R^i(q^i, q^j, z) = P(Q, z)\, q^i$$

where dependence on γ^i, β^i, B^i and E^i is suppressed for exposition. We impose the following conditions:

$$R^i_i \equiv \partial R^i/\partial q^i = P + P_Q\, q^i > 0; \quad R^i_{ii} \equiv \partial^2 R^i/\partial q^{i2} = 2\, P_Q + P_{QQ}\, q^i < 0;$$
$$R^i_j \equiv \partial R^i/\partial q^j = P_Q\, q^i < 0; \quad R^i_{ij} \equiv \partial^2 R^i/\partial q^i\, \partial q^j = P_Q + P_{QQ}\, q^i < 0.$$

In case of pure competition, $P_Q = 0$ and the first inequality follows directly. For other market structures, the first inequality states that marginal revenue is positive. The second inequality is standard: it states that revenues are increasing at a declining rate. The negative cross effects concerning the total and marginal revenues, R^i_j and R^i_{ij}, mean that the rentable space offered by both agents are substitutes.

We further adopt the convention that higher values of z lead to higher total and marginal revenues:

$$R^i_z \equiv \partial R^i/\partial z = P_z\, q^i > 0; \quad R^i_{iz} \equiv \partial^2 R^i/\partial q^i\, \partial z = P_z + P_{Qz}\, q^i > 0.$$

This may be because $P_z > 0$ and $P_{Qz} > 0$.

Agent i also incurs operating costs denoted by $O^i(q^i)$. Consistent with Clark and Kingston (1930)'s study, we assume that operating costs are convex:

$$O^i_i \equiv \partial c^i/\partial q^i > 0; \quad O^i_{ii} \equiv \partial^2 c^i/\partial q^{i2} > 0.$$

Payoff for agent i from being the equity holder

In order to borrow B^i, agent i issues a zero coupon bond with face value D^i. Therefore, the cash flows accrued to agent i in state z are

$$\max \left[R^i(q^i, q^j, z) - O^i(q^i) - D^i, 0 \right].$$

Agent i being the equity holder receives the residual cashflows after incurring operating costs and after debt holders are paid. In the scenario when cash flows are not sufficient to pay the debt holders, debt holders take over the operations of the skyscraper and agent i does not suffer from any penalty. This is the limited liability feature associated with being the equity holder.

For a given debt obligation D^i, height choices q^i and q^j, there is a critical level $\hat{z}^i$ at which agent i is just able to pay her debt. Mathematically, there is a critical level $\hat{z}^i$ such that

$$R^i(q^i, q^j, z) - O^i(q^i) - D^i = 0 \tag{3.2}$$

assuming $\underline{z} \leq \hat{z}^i \leq \overline{z}$. Equation 3.2 implicitly defines $\hat{z}^i$ as a function of q^i, q^j and D^i. As this relationship is important in establishing the principal results of the paper, it is useful to report the following derivatives:

$$d\hat{z}^i/dD^i = 1/R_z^i(\hat{z}^i) > 0; \quad d\hat{z}^i/dD^j = 0;$$
$$d\hat{z}^i/dq^j = -R_i^j(\hat{z}^i)/R_z^i(\hat{z}^i) > 0; \quad d\hat{z}^i/dq^i = -(R_i^i(\hat{z}^i) - O_i^i)/R_z^i(\hat{z}^i).$$

Holding building heights constant, an increase in the face value of debt D^i increases the probability of bankruptcy. Increasing q^j holding q^i and D^i constant decreases expected revenues, which means that agent i will remain solvent in less states of the world. The sensitivity with respect to own building height q^i is unknown, but it forms the basis of the paper. We will show that in the next section that $d\hat{z}^i/dq^i > 0$.

The sensitivity with respect to γ^i, β^i, B^i and E^i is subtle — they affect $\hat{z}^i$ through q^i. The derivatives are

$$d\hat{z}^i/d\gamma^i = \left(d\hat{z}^i/dq^i\right)\left(dq^i/d\gamma^i\right) < 0; \quad d\hat{z}^i/d\beta^i = \left(d\hat{z}^i/dq^i\right)\left(dq^i/d\beta^i\right) < 0;$$
$$d\hat{z}^i/dB^i = \left(d\hat{z}^i/dq^i\right)\left(dq^i/dB^i\right) > 0; \quad d\hat{z}^i/dE^i = \left(d\hat{z}^i/dq^i\right)\left(dq^i/dE^i\right) > 0.$$

Consider for the moment the sensitivity of fixed cost γ^i on $\hat{z}^i$. The negative sign indicates that as the fixed costs of construction increase, probability of default decreases. This is surprising upon first glance. Due to financial constraints in equation 3.1, an increase in γ^i lowers q^i. This leads to lower operating costs O^i and higher revenues R^i (since $R_{ii}^i < 0$). This means that total operating cash flows are higher in every state z. Therefore, holding D^i constant, the agent remains solvent in more states of the world. In this manner, an increase in fixed costs decreases probability of default. Similar interpretations apply to sensitivities with variable costs β^i, and financial investment variables B^i and E^i.

Stage 1   **Stage 2**

- **Each owner i = {1,2}
  chooses the amount of
  debt Dⁱ, and the
  amount of equity Eⁱ**

- **Each owner i = {1,2} chooses
  skyscraper height qⁱ**
- **Both owners choose height
  simultaneously**
- **Random variable z is realized**
- **Rents R(qⁱ,qʲ,z) is received**
- **Debt holders and equity
  holders get paid**
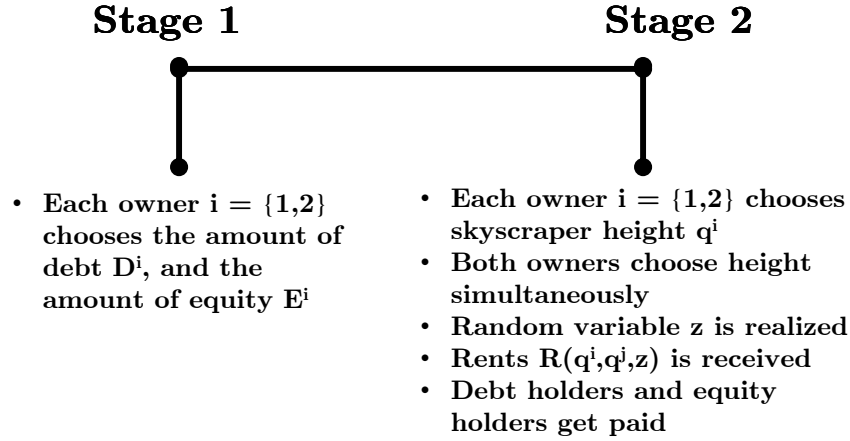
Figure 3.11: *This figure describes the two stages on the model.*

Equilibrium decision stages

Figure 3.11 describes the two decision stages. Financial decisions take place in stage 1. Specifically, each agent chooses the levels of equity E^i and borrowing B^i in this stage. In order to borrow B^i, they issue a zero coupon bond with face value D^i. Risk neutral investors who hold the debt evaluate it at fair market value prior to the resolution of uncertainty. Product market competition takes place in stage 2. Specifically, agent i chooses optimal building height q^i simultaneously with agent j in stage 2 (taking borrowing and equity levels as given). After z is realized, agent i receives cash flows based on equation 3.2. Debt holders either receive the face value of the bond D^i in the event $\{z \geq \hat{z}^i\}$ and they take over the operations in the complementary event $\{z < \hat{z}^i\}$.

The sequential nature reflects two important points. First, choice of building height will depend on the capital structure. Second, rational decision makers (both agents and risk neutral investors of debt) anticipate the influence of debt on building heights, which in turn affects the financial decisions in stage 1. Formally, the stage 2 outcome is a Cournot outcome in building heights which is correctly anticipated by both agents (and investors) while choosing debt levels in stage 1. An agent may prefer to bluff her rival by publicly announcing the construction of a very tall skyscraper. Such an announcement will deter the rival from undertaking her own project. Our equilibrium concept, sub-game perfect equilibrium, rules out such incredible threats.

Given debt levels $\mathbf{D} = (D^i, D^j)$, the expected value to agent i is[2]

$$V^i(q^i, q^j, D^i; B^i, E^i) = \int_{\hat{z}^i}^{\overline{z}} \left(R^i(q^i, q^j, z) - O^i(q^i) - D^i \right) f(z)\, dz - L^i(E^i) - W^i(B^i). \quad (3.3)$$

---

[2]The dependence on γ^i and β^i is suppressed as these are not control variables.
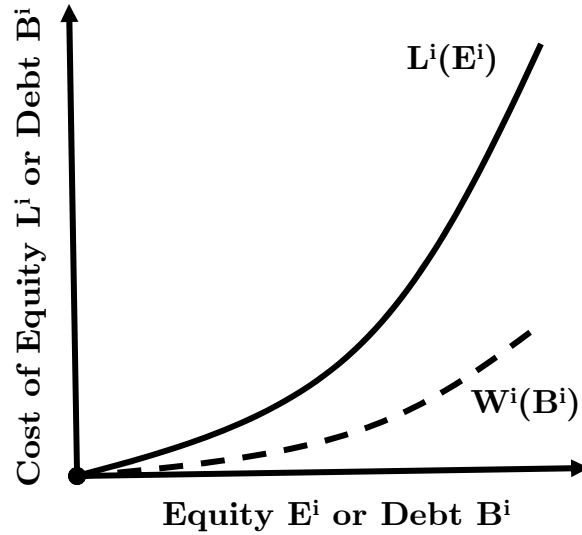
Figure 3.12: *This figure describes the opportunity cost of issuing E^i amount of equity $L^i(E^i)$ and the opportunity cost of issuing B^i amount of debt $W^i(B^i)$.*

The terms $L^i(E^i)$ and $W^i(B^i)$ represent agent i's opportunity cost of equity and debt respectively. The opportunity costs include transaction costs associated with issuing equity and debt. We further assume that opportunity costs of debt and equity are convex:

$$\partial L^i/\partial E^i > 0; \quad \partial W^i/\partial B^i > 0 \quad \partial^2 L^i/\partial E^{i2} > 0; \quad \partial^2 L^i/\partial B^{i2} > 0.$$

Our results will rely on two slight related conditions:

$$\partial L^i/\partial E^i\,(x) \geq \partial W^i/\partial B^i\,(x) \quad \forall\, x$$

and

$$\partial L^i/\partial E^i\,(x_2) - \partial W^i/\partial B^i\,(x_2) \geq \partial L^i/\partial E^i\,(x_1) - \partial W^i/\partial B^i\,(x_1) \text{ for } x_2 > x_1.$$

These two conditions are sufficient when L^i is more convex than W^i. Figure 3.12 depicts $L^i(E^i)$ and $W^i(B^i)$ graphically. We refer to the second inequality as the wedge between opportunity cost of debt and equity. The wedge between opportunity costs of financing are crucial to the analysis and they require further justification. Consider for instance an increase in wealth constrained agent i's equity E^1, which she finances through her own wealth. As a result, agent i is less diversified. Now, if she finances with equity E^2 such that $E^2 > E^1$, then she is even less diversified. Therefore, her opportunity cost of equity is convex. On the other hand, an increase in borrowing B^i does not make the agent less diversified. But, the increase in leverage makes her investment more volatility, which reduces utility. Therefore, her opportunity cost of debt is convex. However, this reduction is utility is less than the the lack of diversification resulting from equity investment. In this manner,

the opportunity cost of equity is more convex than the opportunity cost of debt. Another reason for lower opportunity cost of debt is due to tax. Coupon payments on corporate debt are tax deductible, which benefit the agent.

Because debt market is competitive, the risky asset B^i offers the same expected value as investing B^i in a money market account with risk free rate r. Therefore,

$$B^i \, (1+r) = D^i \, (1 - F(\hat{z}^i)) \, + \, \delta \int_{\underline{z}}^{\hat{z}^i} \left(R^i(q^i, q^j, z) - O^i(q^i) \right) f(z) \, dz \qquad (3.4)$$

In the event $\{z > \hat{z}^i\}$, debt holders are paid off in full, which is reflected in the survival probability $(1 - F(\hat{z}^i))$. In the event $\{z \leq \hat{z}^i\}$, limited liability applies. The agent declares bankruptcy and the debt holders become the residual claimants after incurring bankruptcy costs. Bankruptcy imposes extra costs on bondholders. These costs may include legal fees and transaction costs associated with reorganizing the firm (management of the building) and transferring ownership to the bondholders. We capture these costs in a reduced form. Specifically, we assume that in case of bankruptcy, debt holders recover a fraction $\delta \in [0, 1]$ of the operating cash flows.

Equation 3.4 implicitly links the financial investment $B^i + E^i$, to the face value of the bond D^i. Therefore, in the first stage, equilibrium will ensure that

$$D^i \, = \, D^i(B^i, E^i).$$

The model presented here is the simplest model that captures our interpretation of Raskob's letter. The model captures the effect of debt on the resulting height of the skyscraper (output market) in the presence of oligopolistic competition and competitive capital markets. The analysis of the output market is next.

3.4 Optimal Height choice

This section highlights the influence of competition and limited liability on the optimal height choice.

Equity Value Maximization

Assuming an interior solution, the first order condition for optimal building height is

$$V_i^i(q^i, q^j, D^i; B^i, E^i) = \int_{\hat{z}^i}^{\overline{z}} \left(R_i^i - O_i^i \right) f(z) \, dz \, = \, 0. \qquad (3.5)$$

The first term of eqn 3.5 is the conditional expected marginal revenue received by agent i when q^i is increased. The second term of eqn 3.5 is the conditional expected marginal

operating cost incurred by agent i when q^i is increased. The integral represents the solvent states of the world.

Since $O_i^i > 0$, $\int_{\hat{z}^i}^{\overline{z}} O_i^i f(z) \, dz > 0$. As R_i^i increases in z, it follows that R_i^i evaluated at $\hat{z}^i$ must be negative since the weighted average of $R_i^i - O_i^i$ over $\hat{z}^i$ and strictly better states is zero. Mathematically,

$$R_i^i(\hat{z}^i) - O_i^i < 0.$$

This implies that $d\hat{z}^i/dq^i > 0$ as claimed before. These expressions form the basis of the comparative statics with regards to leverage.

The second order condition is

$$V_{ii}^i = \int_{\hat{z}^i}^{\overline{z}} \left(R_{ii}^i - O_{ii}^i \right) f(z) \, dz \ - \ \left(R_i^i(\hat{z}^i) - O_i^i \right) f(\hat{z}^i) \frac{d\hat{z}^i}{dq^i}.$$

The first term is negative while the second term is positive. Therefore, the second order condition may not hold for all levels of D^i. For the moment, take $D^i = 0$ so that agent does not borrow. This means $\hat{z}^i = \underline{z}$ and the second term drops out. In this case, the second order condition is negative. Now, take the polar opposite case where D^i is such that $\hat{z}^i = \overline{z} - \epsilon$ for $\epsilon > 0$ so that agent is highly leveraged. The first term drops out and the second order condition is positive. Therefore, it seems natural that the second order condition will be satisfied for intermediate values of debt D^i. In addition, we also require a local stability condition. These two assumptions are given below.

ASSUMPTION 10.

$$V_{ii}^i < 0; \quad V_{ii}^i V_{jj}^j - V_{ij}^i V_{ji}^j > 0. \tag{3.6}$$

The second inequality ensures unique equilibrium. A sufficient condition for unique equilibrium is that whenever they intercept, the best response function of agent i is steeper than the best response function of agent j. In turn, a sufficient condition for this to hold is that the slope of the best response function is less than 1 in absolute value — equation 3.6 ensures this restriction.

Equation 3.5 determines the (Cournot) equilibrium heights $\mathbf{q^c} = (q^{ic}, q^{jc})$. To reiterate, these equilibrium heights are dependent on debt levels $\mathbf{D}$, construction costs (γ^i, β^i) for $i \in \{1, 2\}$ and product market competition. Next, we examine the sensitivity of building heights for arbitrary levels of financial investment $B^i + E^i$.

Effect of leverage on building heights

In order to get further insight regarding the effects of leverage, consider the following example. Suppose the inverse demand curve is multiplicative in z:

$$P(Q, z) = z \, p(Q)$$

with $p_Q < 0$. The first order condition is

$$V_i^i = \int_{\hat{z}^i}^{\overline{z}} \left[z\left(p + p_Q\, q^i\right) - O_i^i \right] f(z)\, dz = 0.$$

This equation can be simplified to

$$p\,\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)] - O_i^i\,\mathbb{E}[\mathbb{I}(z \geq \hat{z}^i)] + q^i\, p_Q\,\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)] = 0.$$

where $\mathbb{I}$ is the indicator function. This has a simple interpretation. The first two terms yield profitability from an extra unit of building height, which is equal to the difference between expected price and expected marginal cost. Note that the expected price and the expected marginal cost are evaluated in states of nature in which agent i is solvent. The third term represents the effect of this extra unit of height. Due to an increase in height, expected price decreases by $p_Q\,\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)]$ which affects the q^i units already produced. A more useful explicit form is obtained by dividing the first order condition by $\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)]$, yielding

$$p - O_i^i\, \frac{\mathbb{E}[\mathbb{I}(z \geq \hat{z}^i)]}{\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)]} + q^i\, p_Q = 0$$

where marginal cost is adjusted for leverage and uncertainty. The term $O_i^i\, \frac{\mathbb{E}[\mathbb{I}(z \geq \hat{z}^i)]}{\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)]}$ can be interpreted as modified marginal cost. Denote the modified marginal cost with zero leverage by $\tilde{O}_i^i$ so that

$$\tilde{O}_i^i \equiv O_i^i\, \frac{1}{\mathbb{E}[z]}.$$

Lemma 1 examines how the modified marginal cost changes with an exogenous change in debt by agent i.

Lemma 1. *An increase in debt, D^i, by agent i causes a decrease in the modified marginal cost, i.e.,*

$$O_i^i\, \frac{\mathbb{E}[\mathbb{I}(z \geq \hat{z}^i)]}{\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)]} \leq \tilde{O}_i^i$$

Proof. The proof is mechanical. Note that in the limiting case when agent i does not borrow, the ratio

$$\lim_{D^i \downarrow 0} O_i^i\, \frac{\mathbb{E}[\mathbb{I}(z \geq \hat{z}^i)]}{\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)]} \Big/ \tilde{O}_i^i = \frac{\mathbb{E}[\mathbb{I}(z \geq \hat{z}^i)]}{\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)]} \Big/ \mathbb{E}[z] \Rightarrow 1.$$

If $D^i > 0$, then $\hat{z}^i > \underline{z}$. Then

$$
\begin{aligned}
\frac{d\left[\frac{\mathbb{E}[\mathbb{I}(z \geq \hat{z}^i)]}{\mathbb{E}[z\,\mathbb{I}(z \geq \hat{z}^i)]}\Big/ \mathbb{E}[z]\right]}{d\hat{z}^i} &= \frac{1}{\mathbb{E}[z]} \frac{d\left[\int_{\hat{z}^i}^{\overline{z}} f(z)\,dz \Big/ \int_{\hat{z}^i}^{\overline{z}} z\,f(z)\,dz\right]}{d\hat{z}^i} \\
&= \frac{1}{\mathbb{E}[z]}\left[\frac{-f(\hat{z})}{\int_{\hat{z}^i}^{\overline{z}} z\,f(z)\,dz} + \frac{\int_{\hat{z}^i}^{\overline{z}} f(z)\,dz}{\left(\int_{\hat{z}^i}^{\overline{z}} z\,f(z)\,dz\right)^2}\,\hat{z}^i\,f(\hat{z}^i)\right] \\
&= \frac{1}{\mathbb{E}[z]}\frac{f(\hat{z}^i)}{\int_{\hat{z}^i}^{\overline{z}} z\,f(z)\,dz}\left[\frac{\int_{\hat{z}^i}^{\overline{z}}(\hat{z}^i - z)\,f(z)\,dz}{\int_{\hat{z}^i}^{\overline{z}} z\,f(z)\,dz}\right] \\
&< 0.
\end{aligned}
$$

$\square$

Brander and Lewis [1986] gave an elegant analogy to this setup. They equate the aggressiveness due to leverage to a *Hail Mary* pass in a football game. A football team that is behind in the final seconds of the game often takes chances with a Hail Mary pass. The worst outcome for the team is a loss and it will lose anyway without a change in strategy. So, the team becomes aggressive. In the same manner, an increase in leverage increases the chances of bankruptcy. Since the modified marginal cost decreases, agent increases building height to equate expected marginal revenues and expected marginal costs. Agent i becomes more aggressive. Proposition 1 generalizes the intuition behind the multiplicative example. It is a formal statement of the aggressive strategy followed by agent i and the resulting defensive strategy by agent j due to an increase in debt, D^i.

Proposition 1. (Effect of leverage on heights) *An increase in debt, D^i, by agent i causes her to build taller and her rival, agent j, to build shorter, i.e.*

$$d\,q^i/d\,D^i > 0; \quad d\,q^j/d\,D^i < 0.$$

Proof. Totally differentiating the first order condition, eqn 3.5, we get

$$V_{ii}^i dq^i + V_{ij}^i dq^j + V_{iD^i}^i dD^i = 0$$

and

$$V_{ji}^j dq^i + V_{jj}^j dq^j + V_{jD^j}^j dD^j = 0$$

Writing the two equations in matrix form and using Cramer's rule to invert a 2×2 matrix, we get

$$
\begin{pmatrix} dq^i \\ dq^j \end{pmatrix} = \frac{1}{\triangle} \begin{pmatrix} V_{jj}^j & -V_{ij}^i \\ -V_{ji}^j & V_{ii}^i \end{pmatrix} \begin{pmatrix} -V_{iD^i}^i dD^i \\ -V_{jD^j}^j dD^j \end{pmatrix}
\tag{3.7}
$$

where $\triangle \equiv V_{ii}^i V_{jj}^j - V_{ij}^i V_{ji}^j > 0$.

Solving the equation yields,

$$dq^i/dD^i = -V_{iD^i}^i \, V_{jj}^j \, /\triangle; \quad \text{and} \quad dq^j/dD^i = V_{iD^i}^i \, V_{ji}^j \, /\triangle.$$

Note that

$$V_{ji}^j = \int_{\hat{z}^j}^{\overline{z}} R_{ji}^j \, f(z) \, dz \; - \; \left(R_j^j(\hat{z}^j) - O_j^j \right) \, f(\hat{z}^j) \, d\hat{z}^j /dq^i \; < \; 0$$

Since $V_{jj}^j < 0$ by the second order condition and $V_{ji}^j < 0$ from above, all that is needed is the sign of $V_{iD^i}^i$.

$$V_{iD^i}^i = \left(R_i^i(\hat{z}^i) - O_i^i \right) \, f(\hat{z}^i) \, d\hat{z}^i /dD^i \; < \; 0$$

since $R_i^i(\hat{z}^i) - O_i^i < 0$ from the eqn 3.5. Therefore, it follows that

$$d \, q^i /d \, D^i > 0; \quad d \, q^j /d \, D^i < 0.$$

$\square$

Two corollaries follow directly from Proposition 1.

Corollary 1. *(Effect of leverage on the total supply of rentable space)*

An increase in debt, D^i, by agent i causes total height to increase, i.e.

$$[dq^i + dq^j] \, / \, dD^i \; > \; 0.$$

Proof.
$$[dq^i + dq^j] \, / \, dD^i = V_{iD^i}^i \, (V_{ji}^j - V_{jj}^j) \, /\triangle \; > \; 0 \text{ by equation 3.6.}$$

$\square$

Corollary 2. *(Effect of leverage on the profits of the rival)*

An increase in debt, D^i, by agent i causes the equity value of agent j to decrease, i.e.,

$$dV^j \, / \, dD^i \; < \; 0.$$

Proof. The equity value of agent j is a function of q^i, q^j and D^j. Totally differentiating $V^j(q^i, q^j, D^j)$ with D^i and evaluating the derivative at the optimum yields,

$$dV^j/dD^i \; = \; V_i^j \, dq^i/dD^i \; + \; V_j^j \, dq^j/dD^i \; + V_{D^j}^j \, dD^j/dD^i.$$

$V_j^j = 0$ by the eqn 3.5 for agent j and $dD^j/dD^i = 0$. Since, $V_i^j < 0$, we have that

$$dV^j \, / \, dD^i \; < \; 0.$$

□

The intuition of Proposition 1 and Corollaries 1 and 2 is analogous to a Stackelberg game. The sole reason for debt in the model is to commit to building a taller skyscraper. Formally, this means that the best response function of agent i is shifted out while the best response function of agent j is shifted in due to an increase in D^i. Next, we examine the generality of this result in the next corollary.

Corollary 3. *(Effect of leverage and irrelevance of market structure) An increase in debt, D^i, by agent i causes her to build taller irrespective of the market structure.*

Proof. Based on the proof of Proposition 1, the sign of dq^i/dD^i depends on the sign of $V_{iD^i}^i$, which in turn depends on the sign of $(R_i^i(\hat{z}^i) - O_i^i)$ as

$$V_{iD^i}^i = \left(R_i^i(\hat{z}^i) - O_i^i\right) \, f(\hat{z}^i) \, d\hat{z}^i/dD^i \; < \; 0$$

In case of pure competition, $P_Q(Q, z) = 0$ and equation 3.5 still requires that $(R_i^i(\hat{z}^i) - O_i^i) < 0$.

In case of monopoly, $P_Q(Q, z) < 0$ and $q^j = 0$. But, equation 3.5 still requires that $(R_i^i(\hat{z}^i) - O_i^i) < 0$.   □

An increase in leverage causes agent i to build taller. This result is entirely due to limited liability and does not depend on the underlying market structure.

Effect of construction costs on building heights

This section examines how construction costs (fixed costs γ^i and variable building costs β^i) affect the choices of building heights by the agents. We present the analysis only for the γ^i case, as the other case with β^i is identical.

Imagine that the fixed costs for agent i increase slightly by $d\gamma^i$. Due to the financial restrictions, agent i decreases height since $q_{\gamma^i}^i < 0$. In response agent j increases height. Mathematically, totally differentiating equation 3.5 yields:

$$V_{ii}^i dq^i + V_{ij}^i dq^j = 0$$

Using the fact that $dq^i = q_{\gamma^i}^i \, d\gamma^i$ and substituting in the above equation yields

$$dq^j / d\gamma^i = -V_{ii}^i q_{\gamma^i}^i / V_{ij}^i > q_{\gamma^i}^i \qquad (3.8)$$

The last inequality is due to equation 3.6. While the observation that agent j increases height due to higher fixed costs incurred by agent i is obvious, the magnitude is not. From equation 3.8, an increase in height by agent j is higher than decrease in height by agent i. In turn the industry supply increases. This is summarized in the following proposition.

Proposition 2. *(**Effect of construction costs on heights**) An increase in either fixed costs γ^i or building costs β^i incurred by agent i causes her to build shorter and her rival, agent j, to build taller. Furthermore, industry supply increases, i.e.*

$$dq^i / d\phi^i < 0; \quad dq^j / d\phi^i > 0; \quad [dq^i + dq^j] / d\phi^i > 0; \quad for \ \phi^i \in \{\gamma^i, \beta^i\}.$$

Using the same arguments as that of Corollary 2, we get the following result.

Corollary 4. *(**Effect of construction costs on the profits of the rival**) An increase in either fixed costs γ^i or building costs β^i incurred by agent i causes the equity value of agent j to increase, i.e.,*

$$dV^j / d\phi^i > 0 \quad for \ \phi^i \in \{\gamma^i, \beta^i\}.$$

Effect of product market competition on building heights

This section examines how Cournot competition between the agents affect their optimal choices of building heights. Specifically, we consider two cases. In the first case, both agents act cooperatively. By cooperation, we mean that they maximize joint profits. In the second case, both agents act non-cooperatively. By non-cooperation, we mean that each agent ignores the negative effect of her actions on the rival. This is the case with Cournot competition. Additionally, we also assume symmetry.

In the first case of cooperation, agent i chooses q^i and agent j chooses q^j to maximize joint profits:

$$\int_{\tilde{z}^i}^{\bar{z}} \left(R^i + R^j - O^i - O^j - D^i - D^j \right) f(z) \, dz$$

where $\tilde{z}^i$ is critical state when both debts are paid. The first order condition is

$$\int_{\tilde{z}^i}^{\overline{z}} \left(R_i^i + R_i^j + R_j^j + R_j^i - O_i^i - O_j^j\right) f(z)\, dz \; = \; 0. \tag{3.9}$$

Mathematically, the terms R_j^i and R_i^j account for the negative effect of one agent's actions on the other. This is pivotal in understanding the role of competition.

Denote $\mathbf{q^m} = (\mathbf{q^{im}}, \mathbf{q^{jm}})$ as the solution to equation 3.9 that maximizes joint profits. Recall that the optimal heights when both agents act non-cooperatively (in Cournot competition) are denoted by $\mathbf{q^c} = (\mathbf{q^{ic}}, \mathbf{q^{jc}})$. Then it has to be that

$$R^i(q^{im}, q^{jm}, z) - O^i(q^{im}) \; > \; R^i(q^{ic}, q^{jc}, z) - O^i(q^{ic}),$$

which implies that

$$\tilde{z}^i \; < \; \hat{z}^i.$$

In the cooperation case, agents account for the negative externality that they impose on the rival. Therefore, their profits are higher in every state of nature as compared with the non-cooperation case. In order words, agents remain solvent in more states of the world.

Using symmetry, equation 3.9 yields,

$$\int_{\tilde{z}^i}^{\overline{z}} \left(R_i^i - O_i^i\right) f(z)\, dz \; = \; \int_{\tilde{z}^i}^{\overline{z}} -R_i^j f(z)\, dz \; > \; 0.$$

Since $R_{iz}^i > 0$ and $\tilde{z}^i < \hat{z}^i$, a weighted average of $R^i - O_i^i$ over $\hat{z}^i$ has to be positive (the above integral is a weighted average over $\tilde{z}^i$):

$$\int_{\hat{z}^i}^{\overline{z}} \left(R_i^i - O_i^i\right) f(z)\, dz \; > \; 0. \tag{3.10}$$

From the analysis of typical Cournot and monopoly games, we know that output is lower in the monopoly game as compared with the Cournot game. Next proposition is a formal statement of this result.

Proposition 3. *Both agents build taller in the case when they act non-cooperatively as compared with the case when they act cooperatively, i.e.*

$$q^{im} \; \leq \; q^{ic}.$$

Proof. We show by contradiction. Suppose not, then $q^{im} > q^{ic}$. It is convenient to simplify notation. Denote operating profits in both cases as

$$\pi_i^i(q^{im}, q^{jm}, z) \; \equiv \; R_i^i(q^{im}, q^{jm}, z) - O_i^i(q^{im}); \quad \pi_i^i(q^{ic}, q^{jc}, z) \; \equiv \; R_i^i(q^{ic}, q^{jc}, z) - O_i^i(q^{ic}).$$

From equations 3.10 and 3.5, we have that

$$\int_{\hat{z}^i}^{\overline{z}} \left[\pi_i^i(q^{im}, q^{jm}, z) - \pi_i^i(q^{ic}, q^{jc}, z) \right] f(z) \, dz > 0.$$

Adding and subtracting $\pi_i^i(q^{ic}, q^{jm}, z)$ inside the integral, we have

$$\int_{\hat{z}^i}^{\overline{z}} \left[\pi_i^i(q^{im}, q^{jm}, z) - \pi_i^i(q^{ic}, q^{jm}, z) + \pi_i^i(q^{ic}, q^{jm}, z) - \pi_i^i(q^{ic}, q^{jc}, z) \right] f(z) \, dz > 0$$

Using the Fundamental theorem of calculus, we have

$$\int_{\hat{z}^i}^{\overline{z}} \left[\int_{q^{ic}}^{q^{im}} \pi_{ii}^i(x, q^{jm}, z) \, dx \right] f(z) \, dz + \int_{\hat{z}^i}^{\overline{z}} \left[\int_{q^{jc}}^{q^{jm}} \pi_{ij}^i(q^{ic}, x, z) \, dx \right] f(z) \, dz > 0.$$

Interchanging the integrals and using symmetry, we have that

$$\int_{q^{ic}}^{q^{im}} \left[\int_{\hat{z}^i}^{\overline{z}} \pi_{ii}^i(x, q^{jm}, z) f(z) \, dz \right] dx + \int_{q^{ic}}^{q^{im}} \left[\int_{\hat{z}^i}^{\overline{z}} \pi_{ij}^i(q^{ic}, x, z) f(z) \, dz \right] dx > 0.$$

Since, π_{ii}^i and π_{ij}^i are negative, the inner integrals are negative. If $q^{ic} < q^{im}$, the the LHS is negative and the RHS is positive. Therefore, we have a contradiction. □

To summarize, debt and competition cause overbuilding. This is analogous to the asset substitution theory of Jensen and Meckling (1976). Agent i chooses to build taller as she enjoys limited liability. She declares bankruptcy in the bad states of the world, while earning high rates of return in the good states. This is also evident in the letter. Raskob convinced Kaufmann to proceed with an 80 storey building due to the higher return to equity offered by the taller building.

3.5   Capital Structure Choice

In this section, we describe the determinants of the capital structure choice by each agent. Particularly, in the process of determining the optimal mix of debt and equity financing, agent i considers the effect of leverage on the output market (stage 2). Prior to the analysis of the optimal capital structure, we analyze the marginal effect of debt and equity on height.

Marginal effect of debt and equity

Skyscraper height, q^i, depends on financial investment $B^i + E^i$ through equation 3.1. This is the direct effect. The analysis of stage 2 indicates that q^i also depends on D^i which in turn depends on B^i, E^i. This is the indirect effect. Formally, we have

$$q^i(B^i + E^i, D^i) = q^i(B^i + E^i, D^i(B^i, E^i)) = q^i(B^i, E^i).$$

Total differentiation yields

$$dq^i/dB^i = q^i_{B^i} + (dq^i/dD^i)\,(dD^i/dB^i); \qquad dq^i/dE^i = q^i_{E^i} + (dq^i/dD^i)\,(dD^i/dE^i);$$
$$dq^j/dB^i = (dq^j/dD^i)\,(dD^i/dB^i); \qquad dq^j/dE^i = (dq^j/dD^i)\,(dD^i/dE^i). \qquad (3.11)$$

Similarly, B^i, E^i affects the bankruptcy threshold $\hat{z}^i$ directly and indirectly. Total differentiation yields

$$d\hat{z}^i/dB^i = (d\hat{z}^i/dq^i)\,(dq^i/dB^i) + (d\hat{z}^i/dq^j)\,(dq^j/dB^i) + (d\hat{z}^i/dD^i)\,(dD^i/dB^i);$$
$$d\hat{z}^i/dE^i = (d\hat{z}^i/dq^i)\,(dq^i/dE^i) + (d\hat{z}^i/dq^j)\,(dq^j/dE^i) + (d\hat{z}^i/dD^i)\,(dD^i/dE^i).\,(3.12)$$

The magnitude and sign of these terms depends on cost of debt dD^i/dB^i and dD^i/dE^i. Total differentiation of equation 3.4 with respect to E^i yields

$$(dD^i/dE^i)\,(1-F(\hat{z}^i)) - D^i\,f(\hat{z}^i)\,(d\hat{z}^i/dE^i) + \delta \int_0^{\hat{z}^i} \left[d(R^i(q^i,q^j,z)-O^i(q^i))/dE^i\right]\,f(z)\,dz = 0.$$

A more useful expression is obtained by using equation 3.2, yielding

$$(dD^i/dE^i)\,(1 - F(\hat{z}^i)) - (1-\delta)\left[R^i(q^i,q^j,\hat{z}^i) - O^i(q^i)\right]\,f(\hat{z}^i)\,(d\hat{z}^i/dE^i)$$

$$+ \delta \int_0^{\hat{z}^i} \left[d(R^i(q^i,q^j,z)-O^i(q^i))/dE^i\right]\,f(z)\,dz = 0.$$

An increase in equity on the cost of debt can be decomposed into three effects. The first effect is direct: an increase in equity affects the cost of debt only in the states where agent i remains solvent. The second effect concerns bankruptcy threshold: an increase in equity affects height q^i which in turn affects the bankruptcy threshold. Particularly, a change of bankruptcy threshold by $(d\hat{z}^i/dE^i)$ implies that debt holders lose an amount $(1-\delta)\left[R^i(q^i,q^j,\hat{z}^i)-O^i(q^i)\right]\,f(\hat{z}^i)$ in expectation. The third effect concerns recovery amount upon bankruptcy: an increase in equity affects height q^i which in turn reduces the marginal profits due to equation 3.5.

Substituting equations 3.11 and 3.12 and after slight manipulation, we have

$$dD^i/dE^i = \alpha/\Gamma, \qquad (3.13)$$

where

$$\alpha \equiv \left[(1-\delta)\,f(\hat{z}^i)\,q^i_{E^i}\,(d\hat{z}^i/dq^i)\,\left(R^i(q^i,q^j,\hat{z}^i)-O^i(q^i)\right) - \delta\,q^i_{E^i}\int_0^{\hat{z}^i}(R^i_i - O^i_i)f(z)\,dz\right]$$

and

$$\Gamma = (1-F(\hat{z}^i)) - \left((d\hat{z}^i/dq^i)(dq^i/dD^i) + (d\hat{z}^i/dq^j)(dq^j/dD^i) + (d\hat{z}^i/dD^i)\right)(1-\delta)f(\hat{z}^i)$$

$$\left[R^i(q^i, q^j, \hat{z}^i) - O^i(q^i) \right] + \delta \int_0^{\hat{z}^i} \left((R_i^i - O_i^i)(dq^i/dD^i) + R_j^i(dq^j/dD^i) \right) f(z) \, dz.$$

Note that in case of zero leverage ($D^i = 0$), standard calculations yield $\alpha = 0$ and $\Gamma = 1$. This implies that $dD^i/dE^i = 0$. In case of positive leverage, the numerator α is positive. To see this, note that the first term of α is positive since $d\hat{z}^i/dq^i > 0$ by equation 3.5 and the second term of α is positive since $R_i^i(\hat{z}^i) - O_i^i < 0$ by equation 3.5.

Therefore, the sign of dD^i/dE^i depends on the sign of Γ. We show that $\Gamma < 0$ under normal circumstances which are made precise in the next subsection (we label these circumstances as the Base Case). This is natural. An increase in equity while fixing the borrowed amount B^i, reduces the debt to equity ratio. This lowers the probability of insolvency. Therefore, required yield (also known as credit spread) is lower. This in turn reduces the marginal cost of debt.

Similarly, total differentiation of equation 3.4 with respect to B^i yields

$$dD^i/dB^i = \left[(1 + r) + \alpha \right] / \Gamma. \tag{3.14}$$

If debt were completely risk free, then $dD^i/dB^i = 1 + r$. Comparing equations 3.13 and 3.14, we see that under normal circumstances

$$dD^i/dB^i - dD^i/dE^i = (1 + r) / \Gamma < 0. \tag{3.15}$$

The implication of equation 3.13 is surprising. This is due to the strategic nature of debt. On one hand, an increase in borrowing causes the rival to build shorter (Proposition 1), which increases agent i's expected profits. This effect is positive. On the other hand, an increase in borrowing also increases the chances of insolvency. This effect is negative. Due to limited liability, the agent does not bear bankruptcy costs and as a result the positive effect of borrowing dominates the negative effect. Therefore, at the margin, marginal cost of debt due to borrowing is less than the marginal cost of debt due to an increase in equity.

The use of debt and equity ultimately depend on opportunity costs of debt and equity: functions W^i and L^i. We analyze their impact next.

Marginal effect of debt and equity

Denote the equity value of agent i as a function of borrowing B^i and equity capital E^i by letter $\hat{V}^i$:

$$\hat{V}^i(B^i, E^i) = V^i(q^i(B^i, E^i), q^j(B^i, E^i), D^i(B^i, E^i)) =$$

$$\int_{\hat{z}^i}^{\overline{z}} \left(R^i(q^i(B^i, E^i), q^j(B^i, E^i), z) - O^i(q^i(B^i, E^i)) - D^i(B^i, E^i) \right) f(z) \, dz - L^i E^i - W^i B^i.$$

We use the following notation:

$$\partial \hat{V}^i / \partial B^i \equiv V_{B^i}^i; \quad \partial \hat{V}^i / \partial E^i \equiv V_{E^i}^i; \quad \partial^2 \hat{V}^i / \partial B^{i2} \equiv V_{B^i B^i}^i; \qquad \partial^2 \hat{V}^i / \partial E^{i2} \equiv V_{E^i E^i}^i;$$

$$\partial^2 \hat{V}^i / \partial E^i \partial B^i \equiv V_{B^i E^i}^i; \quad \partial^2 \hat{V}^i / \partial E^i \partial E^j \equiv V_{E^i E^j}^i;$$

$$\partial^2 \hat{V}^i / \partial B^i \partial B^j \equiv V_{B^i B^j}^i; \quad \partial W^i / \partial B^i \equiv W_{B^i}^i.$$

The first order conditions associated with stage 1 maximization are

$$
\begin{aligned}
V_{B^i}^i &= \int_{\hat{z}^i}^{\overline{z}} \left[(R_i^i - O_i^i)(\partial q^i / \partial B^i) + R_j^i (\partial q^j / \partial B^i) - (\partial D^i / \partial B^i) \right] f(z)\, dz \\
&- W_{B^i}^i \le 0
\end{aligned}
\tag{3.16}
$$

$$
\begin{aligned}
V_{E^i}^i &= \int_{\hat{z}^i}^{\overline{z}} \left[(R_i^i - O_i^i)(\partial q^i / \partial E^i) + R_j^i (\partial q^j / \partial E^i) - (\partial D^i / \partial E^i) \right] f(z)\, dz \\
&- L_{E^i}^i \le 0
\end{aligned}
\tag{3.17}
$$

We also assume that certain regularity conditions are satisfied:

ASSUMPTION 11.

$$V_{B^i B^i}^i < 0; \quad V_{E^i E^i}^i < 0; \quad V_{B^i B^i}^i V_{E^i E^i}^i - (V_{B^i E^i}^i)^2 > 0; \qquad \textit{Second order condition}$$

$$V_{B^i B^i}^i V_{B^j B^j}^i - (V_{B^i B^j}^i)^2 > 0; \quad V_{E^i E^i}^i V_{E^j E^j}^i - (V_{E^i E^j}^i)^2 > 0; \qquad \textit{Stability condition}$$

The second order conditions ensure that the principal minors of the Hessian of $\hat{V}^i$ change signs at the optimum. The stability condition ensures that equilibrium is unique.

Using equation 3.5, the first order conditions can be simplified to

$$
\begin{aligned}
V_{B^i}^i &= (\partial D^i / \partial B^i) \int_{\hat{z}^i}^{\overline{z}} \left[R_j^i (\partial q^j / \partial D^i) - 1 \right] f(z)\, dz \\
&- W_{B^i}^i \le 0 \quad (= 0 \text{ if } B^i > 0.)
\end{aligned}
\tag{3.18}
$$

$$
\begin{aligned}
V_{E^i}^i &= (\partial D^i / \partial E^i) \int_{\hat{z}^i}^{\overline{z}} \left[R_j^i (\partial q^j / \partial D^i) - 1 \right] f(z)\, dz \\
&- L_{E^i}^i \le 0 \quad (= 0 \text{ if } E^i > 0.)
\end{aligned}
\tag{3.19}
$$

Now consider the change in profits starting at zero financial investment level: $B^i = 0$ and $E^i = 0$. The first order conditions become

$$V_{B^i}^i = -(1 + r) - W_{B^i}^i [B^i = 0] \le 0 \tag{3.20}$$

$$V_{E^i}^i = -L_{E^i}^i [E^i = 0] \le 0 \tag{3.21}$$

Upon inspection, if the opportunity cost of equity is sufficiently high $L_{E^i}^i[E^i = 0] > (1+r) + W_{B^i}^i[B^i = 0]$, then $V_{B^i}^i > V_{E^i}^i$. This means that marginal increase in profits from borrowing are greater marginal increase in profits from issuing equity. Therefore, agent i will pursue debt financing. The next proposition is a summary of this result.

Proposition 4. (Positive levels of debt) *If the opportunity cost of issuing equity is sufficiently high, so that*

$$L_{E^i}^i|_{E^i=0} > (1+r) + W_{B^i}^i|_{B^i=0},$$

then the skyscraper will be debt financed (at least partially).

Proposition 4 is natural. It makes the simple point that links the benefits of debt and the cost of debt. In the case when the agent is wealth constrained so that the opportunity cost of equity is high, then he will find it beneficial to access capital markets. Even if the agent is not wealth constrained, due tax reasons, opportunity cost of equity may be higher than opportunity cost of debt. We expect this to be the standard case. This matches the anecdotal evidence that most buildings are debt financed.

Next we assume that an internal solution exists. We distinguish two cases concerning $\beta \equiv \int_{\hat{z}^i}^{\overline{z}} \left[R_j^i \left(\partial q^j / \partial D^i \right) - 1 \right] f(z) \, dz$:

- **Case A-** $\beta < 0$ — This is the Base Case.

- **Case B-** $\beta \geq 0$

Consider first Case A. In case of pure competition or monopoly, $R_j^i = 0$ which implies that $\beta < 0$. In case of zero leverage, $(\partial q^j / \partial D^i) = 0$ which also implies that $\beta < 0$. Therefore, it is natural to expect that when either leverage is low or when demand is inelastic, germane environment is characterized by Case A.

Next, consider Case B. Under conditions in which agent i's actions negatively impact the revenues of agent j greatly, germane environment is characterized by Case B.

Substituting equation 3.13 into equation 3.19 yields

$$(\alpha\beta)/\Gamma = L_{E^i}^i(E^*)$$

where E^* is the interior solution of equity. Since $L_{E^i}^i(E^*) > 0$ and $\alpha > 0$, we have to have that

$$\text{sign}(\beta) = \text{sign}(\Gamma).$$

Therefore, in the Base Case where $\beta < 0$, $\Gamma < 0$ as claimed before.

Similarly, interior solution of borrowing in equation 3.18 implies that

$$\partial D^i / \partial B^i = W^i_{B^i}(B^*) / \beta.$$

Comparing equations 3.15, 3.18 and 3.19, in the base case, we have that

$$\partial D^i / \partial B^i |_{(B^*, E^*)} - \partial D^i / \partial E^i |_{(B^*, E^*)} = [W^i_{B^i}(B^*) - L^i_{E^i}(E^*)] / \beta = [1 + r] / \Gamma < 0. \quad (3.22)$$

If $L^i_{E^i}(x) > W^i_{B^i}(x)$, then equation 3.22 implies that

$$B^* > E^*.$$

Equation 3.22 implies that agent i borrows more than he puts up. An immediate testable implication is that the debt to equity ratio of the skyscraper will be higher than 1 or that the debt to asset ratio will b higher than 0.5. This is certainly true in case of the ESB. The debt to asset ratio is 78 % in case of 55-stories and it is 80% in case of 80-stories.

The next two propositions provide the testable implications of the model. The first testable implications concerns overall investment levels and its link with the opportunity cost of equity and debt. The second implication concerns average height and its link with the opportunity cost of equity and debt.

For the first proposition, we assume that

$$\hat{V}^i_{E^i E^i} - \hat{V}^i_{E^i B^i} < 0. \quad (3.23)$$

This is a slightly stronger version of the second order condition in Assumption 11.

Proposition 5. *(Testable Implication 1 — Effect of L^i and W^i on investment)* *Assume opportunity cost of equity is sufficiently higher and steeper than opportunity cost of debt, so that*

$$L^i_{E^i}(x) > W^i_{B^i}(x) \text{ and } L^i_{E^i}(E^i = 0) > W^i_{B^i}(B^i = 0) + (1 + r).$$

Then an opportunity to borrow by an all equity financed agent i results in a net increase in finance investment $B^i + E^i$, if equation 3.23 holds.

Proof. Denote the equity level of an all equity financed agent by E^0. From the first order condition equation 3.18, we know that

$$\hat{V}^i_{E^i}(B^i = 0, E^0) < 0.$$

Now consider arbitrary levels of borrowing B^1 and equity financing E^1 such that

$$B^1 + E^1 = E^0; \text{ and } \hat{V}^i_{B^i}(B^1, E^1) = \hat{V}^i_{E^i}(B^1, E^1).$$

For example, the second equality will be true if $B^1 = B^*$ and $E^1 = E^*$. From the mean value theorem, and equation 3.23, we have that

$$\hat{V}^i_{E^i}(B^1, E^1) - \hat{V}^i_{E^i}(B^i = 0, E^0) = -(\hat{V}^i_{E^i E^i} - \hat{V}^i_{E^i B^i}) B^1 > 0.$$

Therefore, $\hat{V}^i_{B^i}(B^1, E^1) = \hat{V}^i_{E^i}(B^1, E^1) > \hat{V}^i_{E^i}(B^i = 0, E^0)$. Therefore, the agent will increase financial investment. $\square$

Loosely speaking, Proposition 5 implies that when debt is cheap, investment levels in real estate will be higher. We test this by analyzing the number of completed building and the credit spread between AAA and BAA rated securities.

Proposition 6. (Testable Implication 2 — Effect of L^i and W^i on skyscraper height) *Assume both agents are identical. The height of the skyscrapers by both agents increase as wedge between opportunity cost of debt and opportunity cost of equity increases.*

Proof. Assuming an interior solution, we analyze the following:

$$\triangle H \equiv [dq^i + dq^j]/B^i + [dq^i + dq^j]/B^j - [dq^i + dq^j]/E^i - [dq^i + dq^j]/E^j.$$

Direct substitution yields,

$$\triangle H = 2(q^i_{B^i} - q^i_{E^i}) - (\partial q^i/\partial D^i + \partial q^j/\partial D^i)(\partial D^i/\partial B^i - \partial D^i/\partial E^i).$$

The total and partial derivatives are evaluated at (B^*, E^*). If $B^* = E^*$, then the first term is zero and $\triangle H < 0$. As the wedge between opportunity cost of debt and opportunity cost of equity increases, so that

$$L^i_{E^i}(x) - W^i_{B^i}(x)$$

becomes more and more positive, from equation 3.22, we know that

$$B^* >> E^*.$$

That is, the agent borrows more than what he puts up. This makes $(q^i_{B^i} - q^i_{E^i}) > 0$ and $\triangle H$ increases. $\square$

To summarize, we conclude this section by highlighting two major implications of the model. First, cheap credit leads the skyscraper owners to taller than what they would build in autarky with 100% equity financing. Second, cheap credit also leads to an increase in the number of investments.

3.6 Conclusion

Skyscrapers form a distinctive landmark—they stamp an original imprint on the urban landscape. Clusters of skyscrapers that form a city's skyline give a city its distinctive characteristic. Despite the ubiquitous nature of skyscrapers, a general economic theory is scarce. Helsley and Strange (2008) provide a theory behind the ever-increasing height of skyscrapers: they hint at non-economic factors such as ego that lead owners to build taller. We provide an alternate and perhaps a complementary reason. We highlight the role of debt financing that leads owners to build taller.

The theoretical model is based on Brander and Lewis (1986). An increase in skyscraper height implies an increase in supply which decreases prices. In this manner, skyscraper height relates to product market competition. The basic point of the paper is that product market decisions and capital structure decisions are related. With the use of debt, due to the limited liability, equity owner resorts to aggressive behavior. By committing to debt financing, the owner decides to build taller than what she would build if she pursues 100% equity financing.

We justify the commitment role of debt by analyzing a case study. We analyze a letter written by the owner of the Empire State building, John Raskob, to his financier Louis Kauffman. The letter explicitly links debt and skyscraper height. We further corroborate our model by analyzing the time series data of buildings in New York.

The overall point we wish to highlight is that capital structure is important. Debt financing leads to distortions in the product market and our paper is one such example.

Bibliography

Abreu, Maria, Henri L.F. de Groot, and Raymond J.G.M. Florax (2004). "Spatial Patterns of Technology Diffusion: An Empirical Analysis using TFP". In: *Working Paper, Tinbergen Institute.*

Aguerrevere, Felipe L. (2009). "Real Options, Product Market Competition, and Asset Returns". In: *Journal of Finance* 64, pp. 957–983.

Alonso, William (1964). *Location and Land Use.* Cambridge University Press, Cambridge.

Arrow, Kenneth J (1969). "Classificatory Notes on the Production and Transmission of Technological Knowledge". In: *American Economic Review* 59, pp. 29–35.

Back, Kerry and Dirk Paulsen (2009). "Open-Loop Equilibria and Perfect Competition in Option Exercise Games". In: *Review of Financial Studies* 22, pp. 4531–4552.

Baptista, Rui (2000). "Do innovations diffuse faster within geographical clusters?" In: *International Journal of Industrial Organization* 18, pp. 515–535.

Barclay, Michael J and Jr Smith Clifford W (1995). "The Maturity Structure of Corporate Debt". In: *Journal of Finance* 50.2, pp. 609–31.

Barker, David and Tim Loughran (2007). "The Geography of S&P 500 Stock Returns". In: *The Journal of Behavioral Finance* 8, pp. 177–190.

Bena, Jan and Lorenzo Garlappi (2010). "Technological Progress, Industry Rivalry, and Stock Returns". In: *Working Paper,University of British Columbia, Sauder School of Management.*

Berk, Jonathan B., Richard C. Green, and Vasant Naik (1999). "Optimal Investment, Growth Options, and Security Returns". In: *Journal of Finance* 54, pp. 1553–1607.

Black, Fischer and John C. Cox (1976). "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions". In: *The Journal of Finance* 31.2, pp. 351–367.

Bolton, Patrick and David S. Scharfstein (1990). "A Theory of Predation Based on Agency Problems in Financial Contracting". In: *The American Economic Review* 80.1, pp. 93–106.

Brander, James A. and Tracy R. Lewis (1986). "Oligopoly and Financial Structure: The Limited Liability Effect". In: *The American Economic Review* 76.5, pp. 956–970.

— (1988). "Bankruptcy Costs and the Theory of Oligopoly". In: *Canadian Journal of Economics* 21.2, pp. 221–43.

Brueckner, Jan K. (1987). "Chapter 20 The structure of urban equilibria: A unified treatment of the muth-mills model". In: Handbook of Regional and Urban Economics 2. Ed. by Edwin S. Mills, pp. 821 –845.

Bustamante, Maria Cecillia (2011). "Strategic Investment, Industry Concentration and the Cross Section of Returns". In: *Working Paper, London School of Economics and Political Science* 681.

Campello, Murillo (2006). "Debt financing: Does it boost or hurt firm performance in product markets?" In: *Journal of Financial Economics* 82.1, pp. 135 –172.

Carlson, Murray, Adlai Fisher, and Ron Giammarino (2004). "Corporate Investment and Asset Price Dynamics: Implications for the Cross-section of Returns". In: *Journal of Finance* 59, pp. 1540–6261.

Carlson, Murray et al. (2011). "Leaders, Followers, and Risk Dynamics in Industry Equilibrium". In: *Working Paper,University of British Columbia, Sauder School of Management.*

Caselli, Francesco and Wilbur John Coleman (2001). "Cross-Country Technology Diffusion: The Case of Computers". In: *The American Economic Review* 91, pp. 328–335.

Chen, Hui, Yu Xu, and Jun Yang (2012). "Systematic Risk, Debt Maturity and the Term Structure of Credit Spreads". In: *Working Paper, Bank of Canada.*

Chevalier, Judith A. (1995). "Capital Structure and Product-Market Competition: Empirical Evidence from the Supermarket Industry". In: *The American Economic Review* 85.3, pp. 415–435.

Chevalier-Roignant, Benot et al. (2011). "Strategic investment under uncertainty: A synthesis". In: *European Journal of Operational Research* 215.3, pp. 639 –650.

Choi, Jaewon, Dirk Hackbarth, and Josef Zechner (2011). "Granularity of Corporate Debt: Theory and Tests". In: *Working Paper, University of Illinois at Urbana-Champaign.*

Clark, W. C. and J. L. Kingston (1930). *The Skyscraper: A Study in the Economic Height of Modern Office Buildings.* American Institute of Steel Construction: New York.

Clayton, Matthew J. (2009). "Debt, investment, and product market competition: A note on the limited liability effect". In: *Journal of Banking &; Finance* 33.4, pp. 694 –700.

Cochrane, John H. (2011). "Presidential Address: Discount Rates". In: *The Journal of Finance* 66.4, pp. 1047–1108.

Comin, D., M. Dmitriev, and E. Rossi-Hansberg (2012). "The Spatial Diffusion of Technology". In: *Princeton, mimeo.*

Comin, D. and B. Hobijn (2004). "Cross-country technology adoption: Making the theories face the facts". In: *Journal of Monetary Economics* 51, pp. 39 –83.

Comin, Diego and Bart Hobijn (2010). "An Exploration of Technology Diffusion". In: *American Economic Review* 100, pp. 2031–59.

Cooper, Ilan (2006). "Asset Pricing Implications of Nonconvex Adjustment Costs and Irreversibility of Investment". In: *Journal of Finance* 61, pp. 139–170.

Diamond, Douglas W. (1991). "Debt Maturity Structure and Liquidity Risk". In: *The Quarterly Journal of Economics* 106.3, pp. 709–737.

Dougal, Casey, Christopher Parsons, and Sheridan Titman (2012). "Urban Vibrancy and Corporate Growth". In: *Working Paper, University of San Diego.*

Duffie, Darrell and David Lando (2001). "Term Structures of Credit Spreads with Incomplete Accounting Information". In: *Econometrica* 69.3, pp. 633–664.

Eckel, Stefanie et al. (2011). "Measuring the effects of geographical distance on stock market correlation". In: *Journal of Empirical Finance* 18, pp. 237–247.

Estrella, Arturo and Gikas A. Hardouvelis (1991). "The Term Structure as a Predictor of Real Economic Activity". In: *The Journal of Finance* 46.2, pp. 555–576.

Faulkender Michael W., Flannery Mark J. Hankins Kristine Watson and Jason M. Smith (2011). "Cash Flows and Leverage Adjustments". In: *Journal of Financial Economics (JFE), Forthcoming.*

Faulkender, Michael (Apr. 2005). "Hedging or Market Timing? Selecting the Interest Rate Exposure of Corporate Debt". In: *Journal of Finance* 60.2, pp. 931–962.

Faulkender, Michael and Mitchell A. Petersen (2006). "Does the Source of Capital Affect Capital Structure?" In: *Review of Financial Studies* 19.1, pp. 45–79.

Feng, Zhilan, Chinmoy Ghosh, and C. Sirmans (2007). "On the Capital Structure of Real Estate Investment Trusts (REITs)". In: *The Journal of Real Estate Finance and Economics* 34.1, pp. 81–105.

Flannery, Mark J. (1986). "Asymmetric Information and Risky Debt Maturity Choice". In: *The Journal of Finance* 41.1, pp. 19–37.

Flannery, Mark J. and Kasturi P. Rangan (2006). "Partial adjustment toward target capital structures". In: *Journal of Financial Economics* 79.3, pp. 469 –506.

Fong, G. (2006). *Credit Market Handbook,* John Wiley & Sons.

Frank, Murray Z and Vidhan K Goyal (2003). "Testing the pecking order theory of capital structure". In: *Journal of Financial Economics* 67.2, pp. 217 –248.

Fudenberg, Drew and Jean Tirole (1985). "Preemption and Rent Equalization in the Adoption of New Technology". In: *Review of Economic Studies* 52, pp. 383–401.

Goldstein, Robert, Nengjiu Ju, and Hayne Leland (2001). "An EBIT-Based Model of Dynamic Capital Structure". In: *Journal of Business* 74.4, pp. 483–512.

Gomes, Joao, Leonid Kogan, and Lu Zhang (2003). "Equilibrium Cross Section of Returns". In: *Journal of Political Economy* 111, pp. 693–732.

Graham, John R and Campbell R Harvey (2001). "The theory and practice of corporate finance: evidence from the field". In: *Journal of Financial Economics* 60, pp. 187 –243.

Grenadier, Steven R (1996). "The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets". In: *Journal of Finance* 51, pp. 1653–79.

Grenadier, Steven R. (1999). "Information revelation through option exercise". In: *Review of Financial Studies* 12, pp. 95–129.

Grenadier, Steven R. (2002). "Option Exercise Games: An Application to the Equilibrium Investment Strategies of Firms". In: *Review of Financial Studies* 15, pp. 691–721.

Griliches, Zvi (1957). "Hybrid Corn: An Exploration in the Economics of Technological Change". In: *Econometrica* 25, pp. 501–522.

Guedes, Jose and Tim Opler (1996). "The Determinants of the Maturity of Corporate Debt Issues". In: *Journal of Finance* 51.5, pp. 1809–33.

Hall, Bronwyn H. and Beethika Khan (2003). "Adoption of New Technology". In: *NBER Working Papers* 9730.

Helsley, Robert W. and William C. Strange (2008). "A game-theoretic analysis of skyscrapers". In: *Journal of Urban Economics* 64.1, pp. 49 –64.

Hoberg, Gerard and Gordon Phillips (2010). "Real and Financial Industry Booms and Busts". In: *Journal of Finance* 65, pp. 45–86.

Huang, Rongbing and Jay R. Ritter (Mar. 2009). "Testing Theories of Capital Structure and Estimating the Speed of Adjustment". In: *Journal of Financial and Quantitative Analysis* 44, pp. 237–271. ISSN: 1756-6916.

Ivaldi, Marc et al. (2003). "The Economics of Tacit Collusion". In: *Working Paper, Institut d'conomie Industrielle (IDEI), Toulouse.*

Jalilvand, Abolhassan and Robert S Harris (1984). "Corporate Behavior in Adjusting to Capital Structure and Dividend Targets: An Econometric Study". In: *Journal of Finance* 39.1, pp. 127–45.

Jensen, Michael C. and William H. Meckling (1976). "Theory of the firm: Managerial behavior, agency costs and ownership structure". In: *Journal of Financial Economics* 3.4, pp. 305 –360.

Johnson, Shane A. (2003). "Debt Maturity and the Effects of Growth Opportunities and Liquidity Risk on Leverage". In: *Review of Financial Studies* 16.1, pp. 209–236.

Ju, Nengjiu and Hui Ou-Yang (2006). "Capital Structure, Debt Maturity, and Stochastic Interest Rates". In: *Journal of Business* 79.5, pp. 2469–2502.

Julio, Brandon, Woojin Kim, and Michael S. Weisbach (2008). "What Determines the Structure of Corporate Debt Issues?" In:

Karshenas, Massoud and Paul L. Stoneman (1993). "Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies: An Empirical Model". In: *The RAND Journal of Economics* 24, pp. 503–528.

Keller, Wolfgang (2002). "Geographic Localization of International Technology Diffusion". In: *The American Economic Review* 92, pp. 120–142.

Kelley, Maryellen and Susan Helper (1999). "Firm Size And Capabilities, Regional Agglomeration, And The Adoption Of New Technology". In: *Economics of Innovation and New Technology* 8, pp. 79–103.

Kim, In Joon, Krishna Ramaswamy, and Suresh Sundaresan (1993). "Does Default Risk in Coupons Affect the Valuation of Corporate Bonds?: A Contingent Claims Model". In: *Financial Management* 22.3, pp. 117–131.

Kogan, Leonid (2004). "Asset prices and real investment". In: *Journal of Financial Economics* 73, pp. 411–431.

Krugman, Paul. (1991). *Geography and Trade.* Cambridge MA: MIT Press.

Kulatilaka, Nalin and Enrico C. Perotti (1998). "Strategic Growth Options". In: *Management Science* 44.8, pp. 1021–1031.

Leary, Mark T. (2002). "The Dynamics of Capital Structure: An Empirical Analysis of a Partially Observable System". In: *Working Paper, Duke University, Fuqua School of Business.*

Leary, Mark T. and Michael R. Roberts (2005). "Do Firms Rebalance Their Capital Structures?" In: *The Journal of Finance* 60.6, pp. 2575–2619.

Leland, Hayne E (1994). "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure". In: *Journal of Finance* 49.4, pp. 1213–52.

Leland, Hayne E and Klaus Bjerre Toft (1996). "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads". In: *Journal of Finance* 51.3, pp. 987–1019.

Lemmon, Michael L., Michael R. Roberts, and Jaime F. Zender (2008). "Back to the Beginning: Persistence and the Cross-Section of Corporate Capital Structure". In: *The Journal of Finance* 63.4, pp. 1575–1608.

Maksimovic, Vojislav (1988). "Capital Structure in Repeated Oligopolies". In: *RAND Journal of Economics* 19.3, pp. 389–407.

Marcus, Alan J. (1983). "The Bank Capital Decision: A Time Series-Cross Section Analysis". In: *The Journal of Finance* 38.4, pp. 1217–1232.

Marshall, Alfred (1961). *Principles of Economics, 8th edition*. New York: Macmillan Co.

Milbradt, Konstantin and Zhiguo He (2012). "Endogenous liquidity and defaultable bonds". In: *Working Paper, University of Chicago*.

Mills, Edwin S. (1967). "An Aggregative Model of Resource Allocation in a Metropolitan Area". In: *The American Economic Review* 57.2, pp. 197–210.

Muth, Richard (1964). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. University of Chicago Press, Chicago.

Myers, Stewart C. (1977). "Determinants of corporate borrowing". In: *Journal of Financial Economics* 5.2, pp. 147 –175.

No, Angela (2008). "Agglomeration effects in the diffusion of advanced manufacturing technologies". In: *University of Toronto, mimeo*.

Novy-Marx, Robert (2007). "An Equilibrium Model of Investment Under Uncertainty". In: *Review of Financial Studies* 20, pp. 1461–1502.

— (2011). "Operating Leverage". In: *Review of Finance* 15, pp. 103–134.

Ozkan, Aydin (2000). "An Empirical Analysis of Corporate Debt Maturity Structure". In: *European Financial Management*.

Pawlina, Grzegorz and Peter Kort (2006). "Real Options in an Asymmetric Duopoly: Who Benefits from Your Competitive Disadvantage?" In: *Journal of Economics & Management Strategy* 15, pp. 1–35.

Pirinsky, Christo and Qinghai Wang (2006). "Does Corporate Headquarters Location Matter for Stock Returns?" In: *Journal of Finance* 61, pp. 1991–2015.

Rotenberg, JJ and DS Scharfstein (1990). "Shareholder-value maximization and product-market competition". In: *Review of Financial Studies* 3.3, pp. 367–391.

Ryan, Bryce and Neal C Gross (1943). "The diffusion of hybrid seed corn in two Iowa communities". In: *Rural Sociology* 8, pp. 15–24.

Salop, Steven C and David T Scheffman (1983). "Raising Rivals' Costs". In: *American Economic Review* 73, pp. 267–71.

Sarig, Oded and Arthur Warga (1989). "Some Empirical Estimates of the Risk Structure of Interest Rates". In: *The Journal of Finance* 44.5, pp. 1351–1360.

Smit, Han T. J. and L. A. Ankum (1993). "A Real Options and Game-Theoretic Approach to Corporate Investment Strategy under Competition". In: *Financial Management* 22.3, pp. 241–250.

Smith, Adam. (1904). *The Wealth of Nations, Edwin Cannan's edition.* London, Methuen & Co.

Stoneman, Paul (2002). *The economics of Technological Diffusion.* Oxford ; Malden, MA : Blackwell Publishers.

Stoneman, Paul and Myung Joong Kwon (1996). "Technology Adoption and Firm Profitability". In: *Economic Journal* 106, pp. 952–62.

Strebulaev, Ilya A. (2007). "Do Tests of Capital Structure Theory Mean What They Say?" In: *The Journal of Finance* 62.4, pp. 1747–1787.

Tauchen, George (1986). "Finite state markov-chain approximations to univariate and vector autoregressions". In: *Economics Letters* 20.2, pp. 177–181.

Teece, David J (1977). "Technology Transfer by Multinational Firms: The Resource Cost of Transferring Technological Know-how". In: *Economic Journal* 87, pp. 242–61.

Vasicek, Oldrich (1977). "An equilibrium characterization of the term structure". In: *Journal of Financial Economics* 5.2, pp. 177–188.

Welch, Ivo (2012). "A Critique of Recent Quantitative and Deep-Structure Modeling in Capital Structure Research and Beyond". In: *SSRN eLibrary.*

Willis, Carol (1995). *Form Follows Finance: Skyscrapers and Skylines in New York and Chicago.* Princeton Architectural Press.

Wongchoti, Udomsak and Fei Wu (2008). "Provincial co-movement in Chinese stock returns". In: *Applied Financial Economics Letters* 4, pp. 171–176.

Young, H. Peyton (2009). "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning". In: *American Economic Review* 99, pp. 1899–1924.

Zhang, Lu (2005). "The Value Premium". In: *Journal of Finance* 60, pp. 67–103.

Chapter 4

Chapter 1: Appendix

4.1 Proof that $X_P^{21} \in (0, X_2^F)$

Define
$$\xi_2(x) \equiv V_2^L(x) - V_2^F(x)$$
and note that X_P^{21} is the value so that $\xi_2(X_p^{21}) = 0$. I show that $X_P^{21} \in (0, X_2^F)$ by intermediate value theorem. Upon inspection, it is clear that $\xi_2(0)$ is negative and $\xi_2(X_2^F)$ is positive. Therefore, it has to be that $\xi_2(x)$ crosses zero at least once.

4.2 Derivation of the Firm Value

Risk Neutral Measure Derivation

Define random variable L_T by
$$L_T = \frac{d\mathbb{Q}}{d\mathbb{P}} \quad \text{on } \mathcal{F}_T,$$
where $\mathbb{Q}$ is known as the risk-neutral measure. Since $\mathbb{Q} << \mathbb{P}$, i.e. $\mathbb{Q}$ is absolutely continuous with $\mathbb{P}$ on $\mathcal{F}_T$, we also have that $\mathbb{Q} << \mathbb{P}$ on $\mathcal{F}_t$ for all $t \leq T$. We define
$$L_t = \frac{d\mathbb{Q}}{d\mathbb{P}} \quad \text{on } \mathcal{F}_t \quad 0 \leq t \leq T.$$
That is, for every t we have that $L_t \in \mathcal{F}_t$, so L is an adapted process known as the likelihood process. Let the dynamics of L_t be given by
$$dL_t = -L_t \, \lambda \, d\mathbb{W}_t.$$
Then by standard calculations,
$$M_s \;=\; M_t \, e^{-r(s-t)} \frac{L_s}{L_t}.$$

Denote the price of any contingent claim, which is a function $g(X_s)$ by V. Then we have that

$$
\begin{aligned}
M_t V(X_t) &= \mathbb{E}_t\left[\int_t^\infty M_s\, g(X_s)\, ds\right] \\
&= \mathbb{E}_t\left[\int_t^\infty M_t\, e^{-r(s-t)}\frac{L_s}{L_t}\, g(X_s)\, ds\right] \\
V(X_t) &= \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^\infty e^{-r(s-t)}\, g(X_s)\, ds\right]
\end{aligned}
$$

The price of any contingent claim is simply the discounted cash flows (where discounting is done with the risk free rate) in the risk neutral measure. The dynamics of X_t in the risk neutral measure are

$$
\frac{dX_t}{X_t} = (r - \delta)dt + \sigma d\mathbb{W}_t.
$$

$$
\mathbb{E}_t^{\mathbb{Q}}[X_s] = X_t e^{(r-\delta)(s-t)}; \quad \mathbb{E}_t^{\mathbb{Q}}[X_s^2] = X_t^2 e^{(2(r-\delta)+\sigma^2)(s-t)}. \tag{4.1}
$$

Firm value from assets in place

The present value of flow of earnings from assets in place is

$$
\begin{aligned}
W_i &= \mathbb{E}_t^{\mathbb{Q}}\int_t^\infty \pi_i(X_s, \theta_i, \theta_j)\, e^{-r(s-t)}\, ds \\
&= \mathbb{E}_t^{\mathbb{Q}}\int_t^\infty \left[e_{i0}(\theta_i, \theta_j) + e_{i1}(\theta_i, \theta_j)X_s + e_{i2}(\theta_i, \theta_j)X_s^2\right] e^{-r(s-t)}\, ds \\
&= \int_t^\infty e_{i0}(\theta_i, \theta_j)\, e^{-r(s-t)}\, ds + \int_t^\infty e_{i1}(\theta_i, \theta_j)\mathbb{E}_t^{\mathbb{Q}}[X_s]\, e^{-r(s-t)}\, ds \\
&\quad + \int_t^\infty e_{i2}(\theta_i, \theta_j)\mathbb{E}_t^{\mathbb{Q}}[X_s^2]\, e^{-r(s-t)}ds \\
&= \frac{e_{i0}(\theta_i, \theta_j)}{r} + \frac{e_{i1}(\theta_i, \theta_j)X_t}{\delta} + \frac{e_{i2}(\theta_i, \theta_j)X_t^2}{2\delta - r - \sigma^2}.
\end{aligned}
$$

W_i^F, W_i^L and W_i^S can be calculated by substituting appropriate values of the θ_i and θ_j.

Derivation of follower growth option $O_i^F(X_t)$

It is convenient to use the derive the following Lemma. Define $\tau_M = \min\{t : X_t \geq M\}$. Then

LEMMA 8.

$$
\mathbb{E}_t[e^{-r(\tau_M - t)}] = \left(\frac{X_t}{M}\right)^\gamma
$$

where θ is given in the paper.

Proof. Define the random process

$$Y(s) = \int_t^s (r - \delta - 0.5\sigma^2)ds + \int_t^s \sigma d\mathbb{W}^{\mathbb{Q}}.$$

Upon inspection, it is clear that Y_s is normally distributed. Now, consider the process $\{e^{-r(s-t)} e^{\gamma y(s)}\}_{s \geq t}$. For $t \leq \tau_M$, this process is a bounded martingale in the risk neutral measure if the coefficient γ is the positive solution to the equation

$$e^{-r(s-t)}\mathbb{E}[e^{\gamma Y(s)}] = 1.$$

The negative solution makes the equation unbounded. The term in the expectation is simply the moment generating function of $Y(s)$. Standard calculations yield the equation for γ as given in the paper.

By optional sampling theorem, we have that

$$e^{-r(\tau_M - t)}\mathbb{E}[e^{\gamma Y(\tau_M)}] = 1; \quad \text{or} \mathbb{E}[e^{-r(\tau_M - t)}e^{\gamma \log(\frac{M}{X_t})}] = 1.$$

This simplifies to

$$e^{\gamma \log(\frac{M}{X_t})}\mathbb{E}[e^{-r(\tau_M - t)}] = 1.$$

Slight algebra yields the quantity to be proved above.

Note that the proof is quite general. Specifically, the proof and the methodology can be generalized to levy process that do not have any upward jumps. For example, the proof can be easily extended to rare disasters. $\square$

With Lemma 8 at hand, consider the follower firm i value when it follows a trigger strategy. That is, suppose that firm i adopts when the demand level reaches a threshold M. That is

$$M = X_t e^{y(\tau_M)}.$$

The first value from this arbitrary strategy is

$$
\begin{aligned}
O_i^F(X_t, M) &= \mathbb{E}_t^{\mathbb{Q}} \left[\left(\int_{\tau_M}^{\infty} (A_i^F X_s - B_i^F) e^{-r(s-t)} ds \right) - (1-\kappa) I_0 e^{-r(\tau_M - t)} \right] \\
&= \mathbb{E}_t^{\mathbb{Q}} \left[\left(e^{-r(\tau_M - t)} \right) \left(\int_{\tau_M}^{\infty} (A_i^F X_s - B_i^F) e^{-r(s-\tau_M)} ds - (1-\kappa) I_0 \right) \right] \\
&= \mathbb{E}_t^{\mathbb{Q}} \left[\left(e^{-r(\tau_M - t)} \right) \mathbb{E}_{\tau_M}^{\mathbb{Q}} \left[\left(\int_{\tau_M}^{\infty} (A_i^F X_s - B_i^F) e^{-r(s-\tau_M)} ds - (1-\kappa) I_0 \right) \right] \right] \\
&= \mathbb{E}_t^{\mathbb{Q}} \left[\left(e^{-r(\tau_M - t)} \right) \left(\frac{A_i^F X_{\tau_M}}{\delta} - \frac{B_i^F}{r} - (1-\kappa) I_0 \right) \right] \\
&= \mathbb{E}_t^{\mathbb{Q}} \left[e^{-r(\tau_M - t)} \right] \left(\frac{A_i^F M}{\delta} - \frac{B_i^F}{r} - (1-\kappa) I_0 \right) \\
&= \left(\frac{X_t}{M} \right)^{\gamma} \left(\frac{A_i^F M}{\delta} - \frac{B_i^F}{r} - (1-\kappa) I_0 \right).
\end{aligned}
$$

Now, I solve for optimal threshold X_i^F. This can be simply done by differentiating $O_i^F(X_t, M)$ with respect to M and setting that equal to zero. Solving that equation yields the optimal threshold X_i^F. It is also easy to show that second order condition is negative when $X_t = X_i^F$.

Derivation of simultaneous growth option $O_i^S(X_t)$

This proof is identical to the proof of the follower growth option. Simply change A_i^F to A_i^S, B_i^F to B_i^S and set κ equal to zero.

Derivation of Leader option $O_i^L(X_t)$

$$
\begin{aligned}
O_i^L(X_t, X_j^F) &= \mathbb{E}_t^{\mathbb{Q}} \left[\left(\int_{\tau_j^F}^{\infty} (A_i^L X_s - B_i^L) e^{-r(s-t)} ds \right) \right] \\
&= \mathbb{E}_t^{\mathbb{Q}} \left[\left(e^{-r(\tau_j^F - t)} \right) \left(\int_{\tau_j^F}^{\infty} (A_i^L X_s - B_i^L) e^{-r(s-\tau_M)} ds \right) \right] \\
&= \mathbb{E}_t^{\mathbb{Q}} \left[\left(e^{-r(\tau_j^F - t)} \right) \mathbb{E}_{\tau_j^F}^{\mathbb{Q}} \left[\left(\int_{\tau_j^F}^{\infty} (A_i^L X_s - B_i^L) e^{-r(s-\tau_j^F)} ds \right) \right] \right] \\
&= \mathbb{E}_t^{\mathbb{Q}} \left[e^{-r(\tau_j^F - t)} \right] \left(\frac{A_i^L X_j^F}{\delta} - \frac{B_i^L}{r} \right) \\
&= \left(\frac{X_t}{X_j^F} \right)^{\gamma} \left(\frac{A_i^L X_j^F}{\delta} - \frac{B_i^L}{r} \right).
\end{aligned}
$$

4.3 Derivation of the conditions when firm 2 preempts or not

Firm 2 does not attempt to preempt when it has no incentive to be the leader. Formally, this requires that

$$\xi_2(x) \equiv V_2^L(x) - V_2^F(x)$$

is negative for all $x \in (0, X_2^F)$. Also for ease of notation, define $l_2 - l_1 \equiv l$. Therefore, in order to determine the domain of l values where firm 2 does not preempt, we are interested in finding a pair (x^{**}, l^{**}) that satisfies the following system of equations:

$$\xi_2(x^{**}, l^{**}) = 0 \tag{4.2}$$

$$\left. \frac{\partial \xi_2(x, l^{**})}{\partial x} \right|_{x=x^{**}} = 0 \tag{4.3}$$

In other words, we are interested in a point (x^{**}, l^{**}) at which firm 2s leader function is tangent to the follower function. Substituting the firm values and after slight algebra, we have that $\xi_2(x)$ equals

$$\frac{A_2^F - A_2^L}{\delta} x + \frac{B_2^L - B_2^F}{r} - I_0 + \left(\frac{A_2^L X_1^F}{\delta} - \frac{B_2^L}{r} \right) \left(\frac{x}{X_1^F} \right)^\gamma - \left(\frac{A_2^F X_F^F}{\delta} - \frac{B_2^F}{r} \right) \left(\frac{x}{X_2^F} \right)^\gamma \Bigg|_{x=x^{**}} = 0 \tag{4.4}$$

Differentiating with respect to x gives

$$\frac{A_2^F - A_2^L}{\delta} + \left(\frac{A_2^L X_1^F}{\delta} - \frac{B_2^L}{r} \right) \left(\frac{x}{X_1^F} \right)^\gamma \left(\frac{\gamma}{x} \right) - \left(\frac{A_2^F X_F^F}{\delta} - \frac{B_2^F}{r} \right) \left(\frac{x}{X_2^F} \right)^\gamma \left(\frac{\gamma}{x} \right) \Bigg|_{x=x^{**}} = 0 \tag{4.5}$$

Multiplying equation (4.5) by x/γ, and subtracting from equation 4.4 yields

$$\frac{A_2^F - A_2^L}{\delta} x^{**} \frac{\gamma - 1}{\gamma} + \frac{B_2^L - B_2^F}{r} - I_0 = 0.$$

After slight algebra, we obtain:

$$x^{**} = \frac{\gamma}{\gamma - 1} \frac{\frac{B_2^F - B_2^L}{r} + I_0}{\frac{A_2^F - A_2^L}{\delta}} > 0. \tag{4.6}$$

Substituting equation (4.6) in equation (4.4) yields and implicit equation for l^{**}:

$$\frac{B_2^F - B_2^L}{r} + I_0 + \left(\frac{A_2^L X_1^F}{\delta} - \frac{B_2^L}{r} \right) \left(\frac{x}{X_1^F} \right)^\gamma - \left(\frac{A_2^F X_F^F}{\delta} - \frac{B_2^F}{r} \right) \left(\frac{x}{X_2^F} \right)^\gamma \Bigg|_{x=x^{**}} = 0. \tag{4.7}$$

If $l > l^{**}$, then $\xi(x^{**})$ is negative. This means that firm 2 does not prefer to be the follower in the domain $x \in (0, X_2^F)$. Therefore, firm 1 invests at X_1^{LN} without fear of preemption by firm 2.

4.4 Derivation of the conditions when firm 1 simultaneously invests or becomes the leader

Firm 1 simultaneously invests unless the value from simultaneous investment is less than the value from being the leader. Formally, this requires that

$$\zeta_1(x) \equiv V_1^L(x) - V_1^S(x)$$

is negative for all $x \in (0, X_2^F)$. Therefore, in order to determine the domain of l values where firm 1 simultaneously invests with firm 2, we are interested in finding a pair (x^*, l^*) that satisfies the following system of equations:

$$\zeta_1(x^*, l^*) = 0 \tag{4.8}$$

$$\left. \frac{\partial \zeta_1(x, l^*)}{\partial x} \right|_{x=x^*} = 0 \tag{4.9}$$

In other words, we are interested in a point (x^*, l^*) at which firm 1s leader function is tangent to the simultaneous value function. Substituting the firm values and after slight algebra, we have that $\zeta_1(x)$ equals

$$\frac{A_1^S - A_1^L}{\delta} x + \frac{B_1^L - B_1^S}{r} - I_0 + \left(\frac{A_1^L X_2^F}{\delta} - \frac{B_1^L}{r} \right) \left(\frac{x}{X_2^F} \right)^\gamma - \left(\frac{A_1^S X_1^S}{\delta} - \frac{B_1^S}{r} \right) \left(\frac{x}{X_1^S} \right)^\gamma \Bigg|_{x=x^*} = 0 \tag{4.10}$$

Differentiating with respect to x gives

$$\frac{A_1^S - A_1^L}{\delta} + \left(\frac{A_1^L X_2^F}{\delta} - \frac{B_1^L}{r} \right) \left(\frac{x}{X_2^F} \right)^\gamma \left(\frac{\gamma}{x} \right) - \left(\frac{A_1^S X_1^S}{\delta} - \frac{B_1^S}{r} \right) \left(\frac{x}{X_1^S} \right)^\gamma \left(\frac{\gamma}{x} \right) \Bigg|_{x=x^*} = 0 \tag{4.11}$$

Multiplying equation (4.11) by x/γ, and subtracting from equation 4.10 yields

$$\frac{A_1^S - A_1^L}{\delta} x^* \frac{\gamma - 1}{\gamma} + \frac{B_1^L - B_1^S}{r} - I_0 = 0.$$

After slight algebra, we obtain:

$$x^* = \frac{\gamma}{\gamma - 1} \frac{\frac{B_1^S - B_1^L}{r} + I_0}{\frac{A_1^S - A_1^L}{\delta}} > 0. \tag{4.12}$$

Substituting equation (4.12) in equation (4.10) yields and implicit equation for l^*:

$$\frac{B_1^S - B_1^L}{r} + I_0 + \left(\frac{A_1^L X_2^F}{\delta} - \frac{B_1^L}{r} \right) \left(\frac{x}{X_2^F} \right)^\gamma - \left(\frac{A_1^S X_1^S}{\delta} - \frac{B_1^S}{r} \right) \left(\frac{x}{X_1^S} \right)^\gamma \Bigg|_{x=x^*} = 0. \tag{4.13}$$

If $l < l^*$, then $\zeta(x^*)$ is negative. This means that firm 1 does not prefer to be the leader. Therefore, firm 1 invests at X_1^S.

Chapter 5

Chapter 2: Appendix

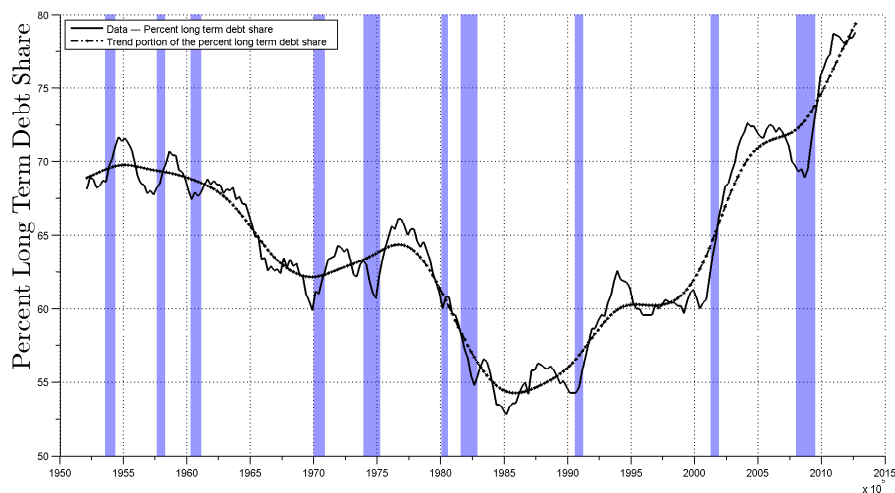## 5.1 Time series plot of the percent long term debt share

Figure 5.1: This is a time series plot of the long term debt share of non-financial corporate business. The solid line the raw data and the dashed dotted line is the trend as calculated by the Hodrick-Prescott filter. The data for the long term debt share is from the Fed Funds flow database (series L.102). The shaded bands in gray are the NBER recession dates.
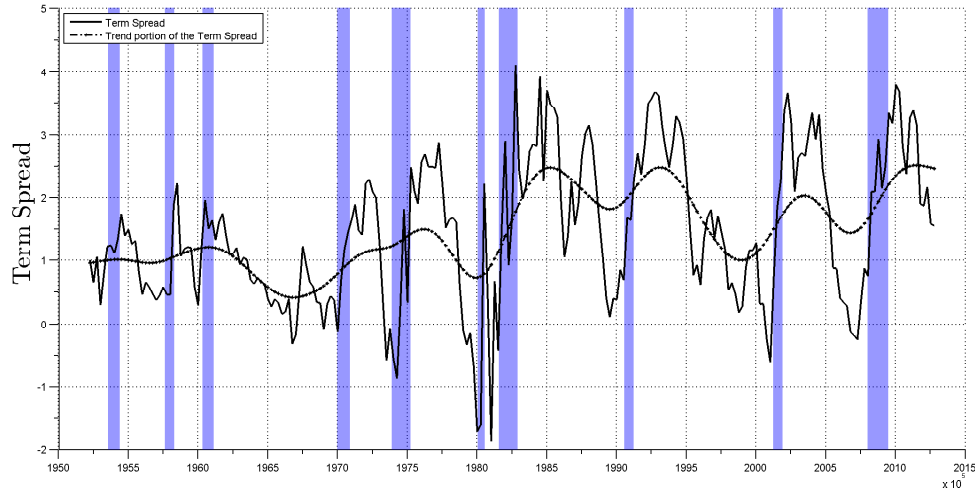
Figure 5.2: This is a time series plot of the term Spread (difference between 10-year Treasure note yield and the 3-month Treasury bill). The solid line is the raw data and the dashed dotted line is the trend as calculated by the Hodrick-Prescott filter. The data is taken from Global Financial database. The shaded bands in gray are the NBER recession dates.

5.2 Time series plot of the term spread

5.3 Derivation of the value of a risky zero coupon bond

$$
\begin{aligned}
D_{\text{zero}}(t, T, r_t; X_0) &= \mathbb{E}_t \left[e^{\int_t^T r_u \, du} \times \{1 - \gamma \, \mathbb{I}(\tau \leq T)\} \right] \\
&= \mathbb{E}_t \left[e^{\int_t^T r_u \, du} \times 1 \right] - \gamma \, \mathbb{E}_t \left[e^{\int_t^T r_u \, du} \times \mathbb{I}(\tau \leq T) \right] \quad (5.1)
\end{aligned}
$$

The first term represents the present value of one dollar upon no default. This expression is simply the value of a default free zero coupon bond $\Lambda(r_t, t, T)$. We use a change of measure to evaluate the second term.

Using Ito's Lemma, the dynamics of the zero coupon bond price are

$$
d\Lambda(r_t, t, T) = \Lambda(r_t, t, T) \left(r_t dt + \sigma_p(t, T) \, d\mathbb{W}_{rt} \right),
$$

where

$$
\sigma_p(t, T) = \frac{\sigma_r \, \Lambda_r(r_t, t, T)}{\Lambda(r_t, t, T)} = \sigma_r \, B(t, T).
$$

Let

$$
\eta_t = e^{\int_0^t \sigma_p(s, T) d\mathbb{W}_{rs} - \int_0^t \sigma_p^2(s, T) ds}.
$$

Note that $\eta_0 = 1$ and $d\eta_t = \eta_t \, \sigma_p(t, T) \, d\mathbb{W}_{rt}$. Furthermore, the since $\sigma_p(t, T)$ is deterministic, the Novikov condition

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T \sigma_p^2(s,T)\,ds}\right] < \infty$$

is satisfied trivially. Therefore, it follows from Girsanov theorem that

$$\left(\begin{array}{c} \mathbb{W}_{vt}^{\mathbb{T}} \\ \mathbb{W}_{rt}^{\mathbb{T}} \end{array}\right) = \left(\begin{array}{c} \mathbb{W}_{vt} \\ \mathbb{W}_{rt} \end{array}\right) - \int_0^t ds \left(\begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array}\right) \left(\begin{array}{c} 0 \\ \sigma_p(s, T) \end{array}\right) \tag{5.2}$$

is a martingale under probability measure $\mathbb{Q}^{\mathbb{T}}$, which is given by

$$\left.\frac{d\mathbb{Q}^{\mathbb{T}}}{d\mathbb{Q}}\right|_{\mathcal{F}_t} = \eta_t \quad \forall \quad t \leq T.$$

Standard calculations yield that

$$\Lambda(r_t, t, T) = \Lambda(r_0, 0, T) \times \eta_t \times e^{\int_0^t r_s ds}, \tag{5.3}$$

and

$$e^{-\int_0^T r_s ds} = \Lambda(r_0, 0, T) \times \eta_T. \tag{5.4}$$

Then, we have that

$$
\begin{aligned}
\mathbb{E}_t\left[e^{-\int_t^T r_u \, du} \times \mathbb{I}(\tau \leq T)\right] &= \mathbb{E}_t\left[e^{-\int_0^T r_u \, du} \times e^{\int_0^t r_u \, du} \times \mathbb{I}(\tau \leq T)\right] \\
&= e^{\int_0^t r_u \, du} \mathbb{E}_t\left[e^{-\int_0^T r_u \, du} \times \mathbb{I}(\tau \leq T)\right] \\
&= e^{\int_0^t r_u \, du} \mathbb{E}_t\left[\Lambda(r_0, 0, T) \times \eta_T \times \mathbb{I}(\tau \leq T)\right] \\
&= e^{\int_0^t r_u \, du} \Lambda(r_0, 0, T) \, \mathbb{E}_t\left[\eta_T \times \mathbb{I}(\tau \leq T)\right] \\
&= e^{\int_0^t r_u \, du} \Lambda(r_0, 0, T) \, \mathbb{E}_t^{\mathbb{T}}\left[\mathbb{I}(\tau \leq T) \times \mathbb{E}_t[\eta_T]\right] \\
&= e^{\int_0^t r_u \, du} \Lambda(r_0, 0, T) \, \eta_t \mathbb{E}_t^{\mathbb{T}}\left[\mathbb{I}(\tau \leq T)\right] \\
&= \Lambda(r_t, t, T) \, \mathbb{E}_t^{\mathbb{T}}\left[\mathbb{I}(\tau \leq T)\right]
\end{aligned}
$$

Next we show that

$$\mathbb{E}_t^{\mathbb{T}}\left[\mathbb{I}(\tau \leq T)\right] = G(T, T, X_0)$$

where

$$G(t; T, X_0) = N\left(\frac{-X_0 - \mu_g(t; T)}{\sqrt{\Sigma(t; T)}}\right) + e^{-\frac{2X_0 \mu_g(t;T)}{\Sigma(t;T)}} N\left(\frac{-X_0 + \mu_g(t; T)}{\sqrt{\Sigma(t; T)}}\right) \tag{5.5}$$

with

$$\mu_g(t; T) = \frac{\int_0^t -\sigma^2(s; T) ds}{2} = \frac{-\Sigma(t; T)}{2};$$

and

$$\Sigma(t;T) = \int_0^t \sigma^2(s;T)ds = \sigma_V^2 t + \frac{\sigma_r^2}{\beta^2}\left(t + e^{-2\beta(T-t)}B_2(t) - 2e^{-\beta(T-t)}B_1(t)\right)$$
$$+ \quad \frac{2\rho\sigma_V\sigma_r}{\beta}\left(t - e^{-\beta(T-t)}B_1(t)\right).$$

and

$$\sigma(t;T) = \sqrt{\sigma_V^2 + \sigma_p^2(t;T) + 2\rho\sigma_V\sigma_p(t;T)}; \ B_1(t) = \frac{(1 - e^{-\beta t})}{\beta}; \ B_2(t) = \frac{(1 - e^{-2\beta t})}{2\beta}.$$

Proof for the expression of $G(t,T,X_0)$

Given the dynamics of X_t, it is well known that the distribution of the first passage hitting times G() in equation (5.5) satisfies the Kolmogorov Backward Equation (KBE). Substituting the dynamics of X_t, KBE could be written as

$$(-\frac{1}{2}\sigma^2(t;T))\frac{\partial G}{\partial X_0} + \frac{1}{2}\sigma^2(t;T)\frac{\partial^2 G}{\partial X_0^2} - \frac{\partial G}{\partial t} = 0$$

with the boundary conditions:

$$G(0;T,X_0) = 0 \text{ for } X_0 > 0 \text{ and } G(t;T,0) = 1.$$

It is sufficient to verify that equation (5.5) satisfies the Kolmogorov Backward Equation. A few tricks are useful. We define the pdf of a standard normal as

$$n(x) = n(-x) = \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}$$

. Straightforward calculations show that

$$n'(x) = -xn(x); \qquad \mu_g(t;T) = \frac{-\Sigma(t;T)}{2}.$$

Next, we evaluate the partial derivatives by brute force. Tedious algebra shows that

$$\frac{\partial G}{\partial X_0} = -n\left(\frac{X_0 - \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right)\frac{1}{\sqrt{\Sigma(t;T)}} - n\left(\frac{X_0 + \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right)\frac{e^{X_0}}{\sqrt{\Sigma(t;T)}} + e^{X_0}N\left(\frac{-X_0 - \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right),$$

$$\frac{\partial G}{\partial t} = n\left(\frac{X_0 - \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right)\left(\frac{X_0\sigma^2(t;T)}{2\Sigma^{3/2}} + \frac{\sigma^2(t;T)}{4\Sigma(t;T)^{1/2}}\right)$$
$$+ \quad e^{X_0}n\left(\frac{X_0 - \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right)\left(\frac{X_0\sigma^2(t;T)}{2\Sigma^{3/2}} + \frac{\sigma^2(t;T)}{4\Sigma(t;T)^{1/2}}\right)$$

and

$$
\begin{aligned}
\frac{\partial^2 G}{\partial X_0^2} &= n\left(\frac{X_0 - \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right)\frac{X_0 - \frac{\Sigma(t;T)}{2}}{\Sigma(t;T)^{\frac{3}{2}}} + n\left(\frac{X_0 + \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right)\frac{(X_0 + \frac{\Sigma(t;T)}{2})e^{X_0}}{\Sigma(t;T)^{\frac{3}{2}}} \\
&\quad - 2n\left(\frac{X_0 + \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right)\frac{e^{X_0}}{\sqrt{\Sigma(t;T)}} + e^{X_0}N\left(\frac{-X_0 - \frac{\Sigma(t;T)}{2}}{\sqrt{\Sigma(t;T)}}\right).
\end{aligned}
$$

Substituting the partial derivatives in the Kolmogorov Backward Equation, we see that equation (5.5) is satisfied.

5.4 Proof of the expression of the bankruptcy cost in a one period debt issuance model

$$
\begin{aligned}
bc(r_0, \zeta, T, V_0) &= \mathbb{E}_0\left[\int_0^T ds\,\gamma\,V_D(r_s, s, T)\,\delta(s - \tau)\,e^{-\int_0^\tau r_u du}\right] \\
&= \mathbb{E}_0\left[\int_0^T ds\,\gamma\,\frac{P_0}{1-\theta}\,\Lambda(s; T, X_0)\,e^{y(T-s)}\,\delta(s - \tau)\,e^{-\int_0^\tau r_u du}\right] \\
&= \mathbb{E}_0\left[\int_0^T ds\,\gamma\,\frac{P_0}{1-\theta}\,\mathbb{E}_s\left[e^{\int_s^T r_u du}\right]\,e^{y(T-s)}\,\delta(s - \tau)\,e^{-\int_0^\tau r_u du}\right] \\
&= \mathbb{E}_0\left[\int_0^T ds\,\gamma\,\frac{P_0}{1-\theta}\,\mathbb{E}_s\left[e^{\int_0^T r_u du}\right]\,e^{y(T-s)}\,\delta(s - \tau)\right] \\
&= \gamma\,\frac{P_0}{1-\theta} \times \int_0^T ds\,e^{y(T-s)}\,\mathbb{E}_0\left[e^{\int_0^T r_u du}\,\delta(s - \tau)\right] \\
&= \gamma\,\frac{P_0}{1-\theta} \times \Lambda(r_0, 0, T)\int_0^T ds\,e^{y(T-s)}\,\mathbb{E}_0\left[\frac{e^{\int_0^T r_u du}}{\Lambda(r_0, 0, T)}\,\delta(s - \tau)\right] \\
&= \gamma\,\frac{P_0}{1-\theta} \times \Lambda(r_0, 0, T)\int_0^T ds\,e^{y(T-s)}\,\mathbb{E}_0^{\mathbb{T}}\left[\delta(s - \tau)\right] \\
&= \gamma\,\frac{P_0}{1-\theta} \times \Lambda(r_0, 0, T)\int_0^T ds\,e^{y(T-s)}\,g(s, T, X_0) \\
&= \gamma\,\frac{P_0}{1-\theta} \times \Lambda(r_0, 0, T)\left[G(T; T, X_0) + \widehat{G}(T; T, X_0)\right].
\end{aligned}
$$

We have used the following property of the dirac delta function. Suppose, we have a random variable $\tilde{x}$, then $\mathbb{E}[\delta(\tilde{x} - x)]$ yields the density at x. To see this, let the density of $\tilde{x}$ be $f(t)$, then $\mathbb{E}[\delta(\tilde{x} - x)] = \int_{-\infty}^{\infty} \delta(\tilde{x} - x)f(t)dt = f(x)$. In our case, $G(t; T, X_0)$ is equal to $Pr(\tau \leq t)$ in the $\mathbb{T}$ forward measure.

$$\mathbb{E}_0^{\mathbb{Q}} \left[\frac{e^{\int_0^T r_u du}}{\Lambda(0; T, X_0)} \delta(s - \tau) \right] = \mathbb{E}_0^{\mathbb{T}} [\delta(s - \tau)] = \Pr(\tau = s) \equiv g(s, T, X_0).$$

5.5   Proof of the expression of the tax benefit in a one period debt issuance model

The market value of debt is given by

$$
\begin{aligned}
L_0 &= \mathbb{E}_0 \left[\int_0^T C e^{-\int_0^s r(u)du} \mathbb{I}_{s<\tau} ds \right] \\
&+ \mathbb{E}_0 \left[P_0 \mathbb{I}_{\tau>T} e^{-\int_0^T r_u du} \right] \\
&+ \mathbb{E}_0 \left[(1-\theta)(1-\gamma) \int_0^T V_D(r_s, s, T) \delta(s - \tau) e^{-\int_0^\tau r_u du} ds \right].
\end{aligned}
$$

The value of debt is composed of three parts: (i) present value of the flow of coupon payments prior to maturity while the firm remains solvent; (ii) present value of principal payment at time T conditional upon not defaulting prior to T and (iii) the present value of the recovery amount conditional upon defaulting at any time before T.

The tax benefit is simply θ times the first expectation which represents the present value of the flow of coupon payments. We evaluate this expression as a difference equation using the fact that debt is issued at par, i.e. $L_0 = P_0$. So,

$$
\begin{aligned}
\theta \times \mathbb{E}_0 \left[\int_0^T C e^{-\int_0^s r(u)du} \mathbb{I}_{s<\tau} ds \right] &= \theta \times P_0 \\
&- \theta \times \underbrace{\mathbb{E}_0 \left[P_0 \mathbb{I}_{\tau>T} e^{-\int_0^T r_u du} \right]}_{\text{Term 1}} \\
&- \theta \times \underbrace{\mathbb{E}_0 \left[(1-\theta)(1-\gamma) \int_0^T V_D(r_s, s, T) \delta(s - \tau) e^{-\int_0^\tau r_u du} ds \right]}_{\text{Term2}}.
\end{aligned}
$$

Note that Term 2 is a $\frac{(1-\gamma)(1-\theta)}{\gamma}$ times $bc(r_0, \zeta, T, V_0)$ and hence we have an expression for it.

Term 1 can be evaluated as follows:

$$
\begin{aligned}
\mathbb{E}_0\left[P_0\mathbb{I}_{\tau>T}e^{-\int_0^T r(u)du}\right] &= \mathbb{E}_0\left[P_0\,\mathbb{I}_{\tau>T}\frac{e^{-\int_0^T r_u du}}{\Lambda(r_0,0,T)}\right]\Lambda(r_0,0,T)\\
&= \mathbb{E}_0\left[\mathbb{I}_{\tau>T}\frac{e^{-\int_0^T r(u)du}}{\Lambda(r_0,0,T)}\right]P_0\,\Lambda(r_0,0,T)\\
&= \mathbb{E}_0^{\mathbb{T}}\left[\mathbb{I}_{\tau>T}\right]P_0\,\Lambda(r_0,0,T)\\
&= P_0\,\Lambda(r_0,0,T)\,(1-G(T,T,X_0)). \quad (5.6)
\end{aligned}
$$

The third equality uses the change of measure formula. The expression for tax benefit follows.

5.6 Derivation of the expression of the coupon C

With the expression for the tax benefit, the coupon rate C can be calculated immediately. We have that Mathematically,

$$
C = \frac{tb(r_0,\zeta,T,V_0)}{\theta\,\mathbb{E}_0\left[\int_0^T e^{-\int_0^s r(u)du}\,\mathbb{I}_{s<\tau}\,ds\right]} = \frac{tb(r_0,\zeta,T,V_0)}{\theta\,\tilde{G}(T,T,X_0)}, \quad (5.7)
$$

where

$$
\tilde{G}(T,T,X_0) = \int_0^T ds\,\Lambda(r_0,0,s)\,(1-G(s,s,X_0)).
$$

Proof of the expression of $\tilde{G}(T,T,X_0)$

$$
\begin{aligned}
\mathbb{E}_0\left[\int_0^T e^{-\int_0^s r(u)du}\,\mathbb{I}_{s<\tau}\,ds\right] &= \int_0^T ds\,\Lambda(r_0,0,s)\,\mathbb{E}_0\left[\frac{e^{-\int_0^s r(u)du}}{\Lambda(r_0,0,s)}\,\mathbb{I}_{s<\tau}\,ds\right]\\
&= \int_0^T ds\,\Lambda(r_0,0,s)\,(1-G(s,s,X_0)\\
&\equiv \tilde{G}(T,T,X_0)
\end{aligned}
$$

The second inequality uses the definition of $G(t,T,X_0)$.

5.7 Expression for the present value of tax benefit at $(N-2)T$ — $\overline{TB_{j;N-2}}$

First note that

$$\mathbb{E}_{N-2}\left[\overline{TB_{j;N-1}} \times \mathbb{I}_{\tau_{N-2}>T}\, e^{\int_{(N-2)T}^{(N-1)T} r_u\, du}\big|r_{(N-1)T} = r_k\right]$$

$$= \mathbb{E}_{N-2}\left[V_{N-1} \times ntb_k \times \mathbb{I}_{\tau_{N-2}>T}\, e^{\int_{(N-2)T}^{(N-1)T} -r_u\, du}\big|r_{(N-1)T} = r_k\right]$$

$$= \mathbb{E}_{N-2}\{V_{N-2}e^{\int_{(N-2)T}^{(N-1)T} (r_u-y-0.5\sigma_v^2 du) + \int_{(N-2)T}^{(N-1)T} \sigma_v dW_{vu} du}$$

$$\times\ ntb_k\, \mathbb{I}_{\tau_{N-2}>T}e^{\int_{(N-2)T}^{(N-1)T} -r_u\, du}\big|r_{(N-1)T} = r_k\}$$

$$= V_{N-2}e^{-yT}\mathbb{E}_{N-2}\left[e^{\int_{(N-2)T}^{(N-1)T} -0.5\sigma_v^2 du + \int_{(N-2)T}^{(N-1)T} \sigma_v dW_{vu} du} \times ntb_k\, \mathbb{I}_{\tau_{N-2}>T}\big|r_{(N-1)T} = r_k\right]$$

$$= V_{N-2}\, e^{-yT}\, ntb_k\, H_{jk}$$

where

$$H_{jk} \doteq H(r_j, r_k, T) = \mathbb{E}_{N-2}\left[e^{\int_{(N-2)T}^{(N-1)T} -0.5\sigma_v^2 du + \int_{(N-2)T}^{(N-1)T} \sigma_v dW_{vu} du} \mathbb{I}_{\tau_{N-2}>T}\big|r_{(N-1)T} = r_k\right].$$

The present value of tax benefits at time $(N-2)T$ are

$$\overline{TB_{j;N-2}} \doteq \overline{TB_{N-2}(r_j, \zeta, T, V_{N-2})}$$

$$= V_{N-2} \times ntb_j + \sum_{k=1}^{M} \pi_{jk}\, \mathbb{E}_{N-2}\left[\overline{TB_{j;N-1}} \times \mathbb{I}_{\tau_{N-2}>T}\, e^{\int_{(N-2)T}^{(N-1)T} -r_u\, du}\big|r_{(N-1)T} = r_k\right]$$

$$= V_{(N-2)T} \times ntb_j + \sum_{k=1}^{M} V_{(N-2)T}\, e^{-yT}\, \pi_{jk}\, ntb_k\, H_{jk}. \tag{5.8}$$

A Closed form approximation of H_{jk}

First, note the expression of H_{jk}.

$$H_{jk} \doteq H(r_j, r_k, T) = \mathbb{E}_{N-2}\left[e^{\int_{(N-2)T}^{(N-1)T} -0.5\sigma_v^2 du + \int_{(N-2)T}^{(N-1)T} \sigma_v dW_{vu} du} \mathbb{I}_{\tau_{N-2}>T}\big|r_{(N-1)T} = r_k\right].$$

This expression involves the sample paths of both the interest rate process and the firm value. Additionally, the expectation is complicated by the fact that the expectation is only derived for paths that start at $r_{(N-2)T} = r_j$ and end up at $r_{(N-1)T} = r_k$. A appropriate way of evaluating this expression involves working with Brownian Bridges for the short rate paths.

We choose a different route. Empirically, the volatility of the firm value is on the order of 20% while the volatility of the interest rate process is on the order of 2%. That is, the firm

value is significantly more volatile than the interest rate process. Therefore, chances are that the default would take place primarily because of the decline in firm value and not due to the changes in the short rate. Consequently, we ignore the conditional expectation involved in H_{jk}. We show that H_{jk} can be well approximated by $\tilde{H}(T, T, X_0)$ where

$$
\begin{aligned}
\tilde{H}(t, T, X_0) &= \mathbb{E}_0 \left[e^{-\sigma_V^2 t/2 + \sigma_V \mathbb{W}_{vt}} \mathbb{I}_{\tau > t} \right] \\
&= N \left(\frac{-X_0 - \mu_h(t, T)}{\sqrt{\Sigma(t, T)}} \right) + e^{-\frac{2X_0 \mu_h(t;T)}{\Sigma(t,T)}} N \left(\frac{-X_0 + \mu_h(t, T)}{\sqrt{\Sigma(t, T)}} \right) \quad (5.9)
\end{aligned}
$$

where

$$
\begin{aligned}
\Sigma(t, T) &= \int_0^t \sigma^2(s; T) ds, \\
&= \sigma_v^2 t + \frac{\sigma_r^2}{\beta^2} \left(t + e^{-2\beta(T-t)} B_2(t) - 2e^{-\beta(T-t)} B_1(t) \right) + \frac{2\rho\sigma_v\sigma_r}{\beta} \left(t - e^{-\beta(T-t)} B_1(t) \right),
\end{aligned}
$$

$$
\mu_h(t, T) = \frac{\int_0^t \sigma^2(s, T) ds}{2} = \frac{\Sigma(t, T)}{2},
$$

$$
\sigma_h(t, T) = \sqrt{\sigma_v^2 + \sigma_p^2(t, T) + 2\rho\sigma_V\sigma_p(t, T)},
$$

and

$$
B_1(t) = \frac{(1 - e^{-\beta t})}{\beta}; \quad B_2(t) = \frac{(1 - e^{-2\beta t})}{2\beta}
$$

The derivation is analogous to that of $G()$. We apply the following steps.

1. Upon inspection, $\tilde{H}(t, T, X_0) = \mathbb{E}\left[e^{-\sigma_v^2 t/2 + \sigma_V \mathbb{W}_{vt}} \mathbb{I}_{\tau > t} \right]$ is already in a change of measure form.

2. After applying the change of measure, we have that

$$
\tilde{H}(t, T, X_0) = \mathbb{E}\left[e^{-\sigma_V^2 t/2 + \sigma_V \mathbb{W}_{vt}} \mathbb{I}_{\tau > t} \right] = \mathbb{E}_0^{\mathbb{V}} \left[\mathbb{I}_{\tau > t} \right].
$$

The new measure $\mathbb{Q}^{\mathbb{V}}$ uses the unlevered firm value as the numeraire.

3.

$$
\mathbb{E}_0^{\mathbb{V}} \left[\mathbb{I}_{\tau > t} \right] = \mathrm{Pr}^{\mathbb{V}}(\tau > t) = 1 - \mathrm{Pr}^{\mathbb{V}}(\tau < t) = 1 - \tilde{H}(t, T, X_0)
$$

As a measure of robustness, we checked the approximation of H_{jk} with $\tilde{H}(t, T, X_0)$ using Monte Carlo simulation. The approximation falls within 5% of the true value for a wide variety of parameter choices. The approximation worked well even when the interest rate process and the firm value are correlated.