AN INTERFACE OF OPERATIONS AND FINANCE

by

DANKO TURCIC

Submitted in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy

Dissertation Advisers: Professors Peter H. Ritchken and Matthew J. Sobel

DEPARTMENT OF OPERATIONS
CASE WESTERN RESERVE UNIVERSITY
AUGUST 2008
CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF GRADUATE STUDIES

We hereby approve the thesis/dissertation of

__________________________
Danko Turcic

candidate for the Doctor of Philosophy degree.*

(signed) ____________________ Matthew J. Sobel
(Chair of the Committee)

__________________________
Peter H. Ritchken

__________________________
Kamlesh Mathur

__________________________
James C. Alexander

(date) ______________________
June 6, 2008

*We also certify that written approval has been obtained for
any proprietary material contained therein.
# Contents

List of Figures ......................................................... vi

Acknowledgements ..................................................... viii

Abstract ................................................................. ix

1 Coordination of Inventory, Capital, and Dividends in Nascent Firms 1

1.1 Introduction ......................................................... 1

1.2 Related Research ................................................... 5

1.3 The Model ............................................................. 7

1.4 Preliminary Results ............................................... 12

1.5 A Myopic Optimum ............................................... 13

1.6 Firm Size and the Worth of Coordination ....................... 18

1.7 Financial Distress .................................................. 22

1.8 Conclusions and Generalizations ................................ 26

2 Valuing Adaptive Quality and Shutdown Options in Production Systems ......................................................... 28
<table>
<thead>
<tr>
<th>Section</th>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>28</td>
</tr>
<tr>
<td>2.2</td>
<td>Basic Model</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>2.2.1 Analysis</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>2.2.2 An Optimal Policy</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>2.2.3 Valuation</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>2.2.4 Quality Option Contribution</td>
<td>42</td>
</tr>
<tr>
<td>2.3</td>
<td>Model Extensions</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>2.3.1 Analysis</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>2.3.2 A Monotone Optimal Policy</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>2.3.3 An Example</td>
<td>54</td>
</tr>
<tr>
<td>2.4</td>
<td>Conclusion</td>
<td>58</td>
</tr>
<tr>
<td>3</td>
<td>Risk Averse Supply Chain Coordination</td>
<td>60</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>60</td>
</tr>
<tr>
<td>3.2</td>
<td>The Model</td>
<td>63</td>
</tr>
<tr>
<td>3.3</td>
<td>The Game</td>
<td>65</td>
</tr>
<tr>
<td>3.4</td>
<td>Analysis</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>3.4.1 A Risk Neutral Case</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>3.4.2 A Risk Averse Case</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>3.4.3 An Example</td>
<td>74</td>
</tr>
<tr>
<td>3.5</td>
<td>Conclusion</td>
<td>76</td>
</tr>
<tr>
<td>A</td>
<td>Appendix to Chapter 1</td>
<td>78</td>
</tr>
<tr>
<td>A.1</td>
<td>The Complete Solution</td>
<td>78</td>
</tr>
</tbody>
</table>
List of Figures

2.1 The market value of the sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$ viewed from time $t_0$, $V_0(S_0)$, when the effort can adapt to market conditions, the cost of input $I$ is constant and $K + I < c$ with parameters $S_0 = $50, $q_s = 0.02$, $I = $25, $K = $5, $c = $40, $r = 0.05$, $\sigma_S = 0.45$, $t_1 - t_0 = 0.25$ years, and $\Delta t = 0.25$ years. 42

2.2 The market value of the sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$ viewed from time $t_0$, $V_0(S_0, I_0)$, when the effort can adapt to market conditions, the cost of input $I_i$ is stochastic and $X < c$ with parameters $S_0 = $50, $q_S = 0.02$, $I_0 = $25, $q_I = 0.01$, $K = $5, $c = $40, $r = 0.05$, $\sigma_S = 0.45$, $\sigma_I = 0.35$, $\rho = 0.2$, $t_1 - t_0 = 0.25$ years, and $\Delta t = 0.25$ years. 43

2.3 The relative value of the quality option $R^{(1)}(S_0; c, \sigma_S|k_f')$ with parameters $S_0 = $50, $q_S = 0.02$, $I = $25, $K = $5, $r = 0.05$, $t_1 - t_0 = 0.25$ years, and $\Delta t = 0.25$ years. 46
2.4 Bounds $k_L^{(i)}$ and $k_U^{(i)}$ defined in Proposition 13 with parameters $S_0 = \$50$, $q_S = 0.02$, $I = \$25$, $K = \$5$, $c = \$100$, $r = 0.05$, $\sigma_S = 0.45$, $t_1 - t_0 = 0.25$ years, $\Delta t = 0.25$ years, $i = 1, 2, \ldots, n$, and $n = 20$.

2.5 Optimal policy at times $t_1$ and $t_2$ as a function of the market price of output $S$ given $q_S = 0.02$, $I = \$2$, $r = 0.05$, $\sigma_S = 0.45$, $\Delta t = 0.25$ years. There are three states: $k^{(1)} = 0$, $k^{(2)} = 0.5$, $k^{(3)} = 1$. The transition probabilities are: $p(0, 0) = 1$, $p(0.5, 0.5) = 0.625$, $p(0.5, 0) = 1 - p(0.5, 0.5)$, $p(1, 1) = 0.75$, $p(1, 0.5) = 1 - p(1, 1)$, and zero otherwise. The costs are: $g(0, 0) = \$0$, $g(0, 0.5) = \$5.25$, $g(0, 1) = \$22.375$, $g(0.5, 0.5) = \$2.75$, $g(0.5, 1) = \$18.625$, $g(1, 1) = \$2.5$, and infinity otherwise.

2.6 Relative value of the quality option at time $t_0$ given $S_0 = \$10$, $k_0 = 0.5$, $k_f' = 1$, $q_S = 0.02$, $I = \$2$, $r = 0.05$, $\sigma_S = 0.45$, $\Delta t = 0.25$ years, and $t_0 = t_1$. There are three states: $k^{(1)} = 0$, $k^{(2)} = 0.5$, $k^{(3)} = 1$. The transition probabilities are: $p(0, 0) = 1$, $p(0.5, 0.5) = 0.625$, $p(0.5, 0) = 1 - p(0.5, 0.5)$, $p(1, 1) = 0.75$, $p(1, 0.5) = 1 - p(1, 1)$, and zero otherwise. The costs are: $g(0, 0) = \$0$, $g(0, 0.5) = \$5.25$, $g(0, 1) = \$22.375$, $g(0.5, 0.5) = \$2.75$, $g(0.5, 1) = \$18.625$, $g(1, 1) = \$2.5$, and infinity otherwise. Note that $\phi_{fix}(S_0, k_0 | k_f') = 0.12$, $V_{fix}(S_0, k_0 | k_f') = 0.018$, $\phi_0(S_0, k_0) = 2.19$, and $V_0(S_0, k_0) = 0.746$. 

vii
This research was conducted over the three-year period from June 2005 to April 2008. I am indebted to Matthew J. Sobel, my principal advisor, and to Peter H. Ritchken for their mentorship; this work simply could not have been completed without their help.

In addition, I would like to thank seminar participants at the Amos Tuck School of Business Administration at Dartmouth College, the John M Olin School of Business, the Kellogg School of Management, and the USC Marshall School of Business for their comments, which helped refine the contents of Chapter 2.

A special thanks goes to Anna L. Leavey for her love and seemingly infinite patience.
This dissertation has three essays. The first one, “Coordination of Inventory, Capital and Dividends in Nascent Firms,” considers how best to coordinate borrowing and inventory decisions in a growing firm. The firm in the model has a single product for which there is an autoregressive demand process, and it optimizes the expected present value of dividends. The results include a myopic policy with coordinated base-stock levels for physical inventory and retained earnings. These levels change with the size of the firm, and we compare them with their counterparts when the firm decentralizes inventory and financial decisions. We find repeatedly that the relative improvement of coordination is greater for smaller firm than for large.

The second essay, “Valuing Adaptive Quality and Shutdown Options in Production Systems,” is concerned with the value of operational flexibility and how best to exploit it. We consider two forms of operational flexibility in a production setting: the option to shut down production temporarily in order to avoid losses, and the opt-
tion to control expected production yield through costly effort. The former is referred to as the *temporary shutdown option*, the latter as the *quality option*. The absolute values of both options are sensitive to costs and prices of finished goods, and they increase with price uncertainty. Moreover, we find that their relative contribution to the value of production technology may increase or decrease as price uncertainty increases; sufficient conditions are provided for both. This is useful to understand when considering an investment in a costly yield management technology.

The research for the third essay, “Risk Averse Supply Chain Coordination,” solves a supply chain bargaining game that is different from the Stackelberg game commonly analyzed in the supply chain contracting literature. The solution, for the case when both the supplier and the retailer are risk neutral, agrees with the solution of the Stackelberg game in terms of the optimal order quantity and it sharpens the existing results in terms of the division of total profit. The solution is also amenable to incorporating risk aversion; we present the risk neutral and risk averse cases and contrast the two.
Chapter 1

Coordination of Inventory, Capital, and Dividends in Nascent Firms

1.1 Introduction

Poor management of inventory and credit are often the culprits when nascent firms struggle (e.g. Ames (1983)). The U.S. Small Business Administration emphasizes the importance of successfully managing financial resources in newly created firms, and observes (sba.gov (2007)) that a nascent firm can improve its financial base either by growing gradually and allowing profits to fund additional growth, or by seeking outside funds such as short-term loans or lines of credit. However, it is unclear whether firms should borrow early on, or whether they should delay borrowing while they build equity through retained earnings. In addition to managing access to capital and liquidity, firms must contain their costs. Effective inventory management is an important element of cost containment, particularly in retail and manufacturing.
firms, but the commonly utilized inventory models are divorced from financial considerations. The models implicitly assume that the theorem of Modigliani and Miller (1958) is valid, although the premises of this theorem are frequently invalid due to factors such as incipient bankruptcy costs.

This paper considers how best to coordinate borrowing and inventory management decisions in recently formed corporate firms. Leveraged firms are extensively analyzed in the finance literature. Merton (1974) considers a static problem of pricing corporate debt when there is a positive probability that the firm will be unable to satisfy some of its debt obligations. Although that research has been extended to multi-period formulations by Geske (1977) and others, the results offer little guidance for non-financial decisions. The Black-Scholes framework introduced in the Merton model is well suited for valuation and hedging, and these are foci of corporate finance, but it is not well suited for optimizing production and related decisions. Although the valuation problem is well understood, there has been little analysis of optimal operating policies when firms have a significant probability of default. Two exceptions are Xu and Birge (2004) and Li et al. (1997) who study the coordination of operating and financial decisions, but the assumptions in both works preclude firm growth. Here we consider the coordination of operating and financial decisions when changes in demand may occur due to growth, seasonality, or other factors.

Answers to questions such as ours are driven by model details. Since we seek to obtain as much insight into coordination as possible, we use simple operational and financial structures in the model and an autoregressive demand process. This type of process is nonstationary but tractable. On the operational side, the firm produces a
single product at a single location for a market with an exogeneously fixed price. The
firm has 100% yields, there is no delay in the availability of goods to satisfy demand,
and it experiences linear costs of production and storage. On the financial side,
the firm can augment its working capital only with short-term borrowing and capital
subscriptions. At first we model financial distress with the costs of reorganization that
follow bankruptcy. This assumption is based on data that show that the majority
of defaults do not result in liquidation (Franks and Torrous (1994)). Instead, the
defaulting firm is often restructured and creditors are compensated with a new debt
issue. Later we consider the opposite extreme, namely dissolution of the firm following
bankruptcy.

The most important measure of the value of an investor-owned firm is the ex-
pected present value of the sequence of its dividends (cf. Brealey and Myers (2006),
Cochrane (2001), and Ross et al. (2006)). So the optimization criterion in this paper
is the expected present value of the dividends net of capital subscriptions. Capital
subscriptions (i.e., additional capital raised from the current investors) occur often in
nascent firms.

We make no attempt to assign a market value to the firm in the model and, as a
consequence, we do not specify a pricing measure. Our discount factor is a placeholder
for what would otherwise be a risk free discount factor under the risk neutral pricing
measure (or an expected value of the pricing kernel under the real-world measure).
This shortcut does not affect the conclusions but it streamlines the exposition.

The results include the complete characterization of an optimal policy that is my-
opic and has base-stock levels for physical goods inventory and retained earnings. The
physical inventory base-stock level is lower than that in the corresponding standard inventory model divorced from financial considerations. The optimal short-term loan follows the well-known pecking-order hypothesis. We compare the effects of optimal coordinated decisions with the effects of decentralized decisions. We find repeatedly that the relative improvement of coordination is greater for small firms than for large. That is, the application of standard inventory and production models may be particularly ill-advised in small firms.

Our proxy for firm size is the deterministic part of the autoregressive demand process, $\mu$. Let $r$ denote the exogeneous market price for the firm’s good. The inventory base-stock level equals $\mu$ plus a factor, and the one for retained earnings equals $-r\mu$ plus a factor. The size of the resulting short-term loan is $r\mu$ plus a factor. None of these factors depend on $\mu$, and the optimal policy induces dividends that grow with firm size at the rate of $r(1 - \rho)$ where $\rho$ is the interest rate for short-term borrowing.

It is important to examine the dependence of the conclusions on the assumptions. For example, bankruptcy is a complex financial and legal process, and in most of the paper the firm’s inability to repay a loan leads to a default penalty, namely the costs of restructuring debt. However, Section 1.7 explains why the replacement of the restructuring assumption with irrevocable dissolution of the firm would preserve key results. Although we introduce a general autoregresssive demand process in Section 1.3, thereafter we employ a first-order process for expository simplicity. Most of the paper’s results would be preserved with a higher-order process. In Section 1.5 we introduce piece-wise linear expressions for the default penalty and the sales revenue
net of inventory-related costs. Most of the qualitative results would be preserved with more general functions but simple explicit formulas would be lost.

The paper is organized as follows. The model is specified in Section 1.3, and leads in Section 1.4 to properties of a relevant dynamic program and the structure of an optimal policy. A myopic policy is shown to be optimal in Section 1.5, and the piece-wise linearity assumption yields simple explicit formulas. Some consequences of firm size are analyzed in Section 1.6 which also compares the myopic optimal policy with a decentralized policy. In Section 1.7 the myopic policy is shown to induce a probability of default that is constant with respect to firm size, and the model is changed to examine the consequences if bankruptcy causes irrevocable dissolution of the firm. The results are summarized in Section 1.8 which briefly discusses the effects of altering the assumptions and model details.

### 1.2 Related Research

There are several sub-literatures on the coordination of operations and finance. They include operational decisions in the presence of foreign exchange exposure (e.g. Kogut and Kulatilaka (1994), Huchzermeier and Cohen (1996), Dasu and Li (1997), Aytekin and Birge (2004), Dong et al. (2006)) and capacity-expansion problems with financial constraints (e.g. Birge (2000), Van Mieghem (2003), and Babich and Sobel (2004)). Some of that research uses approaches similar to ours, but this paper contributes to the sub-literature on the coordination of production and financing decisions and is closest to Li et al. (1997). Their model is the special case of ours with a $0^{th}$-
order autoregressive demand process, namely independent and identically distributed random demands. Also, excess demand is backordered in their paper and lost in ours. The authors find that a myopic policy is optimal and study its properties.

Other papers based on Li et al. (1997) are Hu and Sobel (2005) and Brunet and Babich (2006). The former examines the impact of a firm’s capital structure on its short-term operating and financial decisions. The latter studies the value of trade credit when the firm’s borrowing capacity is uncertain.

Earlier contributions to the coordination of production and financial decisions include Archibald et al. (2004) and Xu and Birge (2004). The former optimizes the probability of financial survival of a start-up firm that manages inventory, but does not seek outside funds. The analysis stems from the hypothesis that nascent firms are more concerned with long-term survival than with profitability. The latter uses a one-period newsvendor framework to show that the firm’s optimal production quantity is a decreasing function of its financial leverage and that misidentifying the company’s optimal leverage ratio decreases the firm’s market value.

The important difference between these models and ours is that they preclude the possibility of firm growth because they either assume identically distributed demands or they are one-period models. A more general demand model is needed here because growth is the chief motivation of the nascent firm that we analyze. Our use of a non-stationary demand process leads to insights that cannot be elicited from models with stationary or static demands.
1.3 The Model

We use a discrete-time multi-period model of a single-product single-location corporate firm that decides at the beginning of each period how much money to borrow, $b_n$, how many units to produce or procure, $z_n$, and how much of a dividend to issue, $v_n$. Then a stochastic demand, $D_n$, occurs and various revenues are received and costs paid.

At the beginning of each period $n$, ($n = 1, 2, ...$) the firm observes the amount of retained earnings, $w_n$ (unconstrained in sign), the current inventory level, $x_n \geq 0$ (implying that excess demand is lost), and the amount of demand in the previous period, $D_{n-1}$. If $w_n < 0$, then the firm was unable to meet some of its debt obligations in period $n-1$ and it is assumed to be in bankruptcy at the start of period $n$. When in bankruptcy, the firm is assessed a default penalty $p(w_n)$, but it does not cease operations. Actual bankruptcy processes can be quite complex. Therefore, for modeling purposes, we make the simplifying assumption of debt refinancing. In the model, the restructuring costs are represented by the default penalty $p(w_n)$. In Section 1.7, we replace restructuring with the complete cessation of operations.

Prior to observing demand in period $n$, the firm makes these decisions: $z_n \geq 0$, $b_n \geq 0$, and $v_n$ (unconstrained in sign). We interpret negative values of $v_n$ as capital subscriptions which occur often among entrepreneurial ventures.

Production and borrowing costs are $cz_n$ and $\rho b_n$, respectively, where $c$ and $\rho$ denote the unit costs of production and borrowing. We assume that the production lead time is negligible; so $x_n + z_n$ is the supply of goods that are available to satisfy the demand
in period \( n \). Borrowing is short-term; that is, the outstanding principal, \( b_n \), is due to be repaid at the end of period \( n \). We assume that the borrowing cost rate, \( \rho \), is a function of the default probability, but that probability turns out to be constant (cf. Section 1.7). So we interpret \( \rho \) as the value taken by the function at the constant default probability.

We assume that the sales revenue net of inventory-related costs, denoted \( g(x_n + z_n, D_n) \), depends on the total supply and demand in period \( n \).

The flow of goods and dollars in the model is subject to conservation constraints

\[
w_{n+1} = w_n - p(w_n) - v_n - cz_n + g(x_n + z_n, D_n) - \rho b_n, \tag{1.1}
\]

\[
x_{n+1} = (x_n + z_n - D_n)^+, \tag{1.2}
\]

a liquidity constraint that prevents the expenditures in period \( n \) from exceeding the sum of retained earnings plus the loan proceeds

\[
w_n + (1 - \rho)b_n \geq cz_n + v_n + p(w_n), \tag{1.3}
\]

and logical constraints

\[
b_n \geq 0 \quad \text{and} \quad z_n \geq 0. \tag{1.3}
\]

By the end of period \( n \), the firm will have observed demand \( D_n \), realized revenue net of inventory costs \( g(x_n + z_n, D_n) \), and repaid the entire loan principal \( b_n \) if \( w_{n+1} \geq 0 \), where \( w_{n+1} \) is given by (1.1). Otherwise, there is a delay in repayment and the default penalty \( p(w_{n+1}) \) is levied.
Demand Model

An autoregressive processes makes it possible for firm size to change over time. These processes are attractive too because they are well understood (cf. Greene (2003), Hamilton (1994)), relatively easily estimated, and encompass many forms of systematic non-stationarity, such as seasonality and trends. Therefore, let

\[ D_n = \sum_{k=1}^{K} \theta_k D_{n-k} + \varepsilon_n \]  

(1.4)

where \( \{\theta_k\} \) are known scalars, and \( \{\varepsilon_n\} \) are independent, identically distributed, and nonnegative random variables.

Let \( \varepsilon \) be a random variable with the same distribution as \( \varepsilon_1 \), and let \( F(\cdot) \) denote the distribution function of \( \varepsilon \). We say that \( q^* \) solves \( F(q^*) = \alpha \) if \( q^* = \sup\{q : F(q) \leq \alpha\} \), and we write \( q^* = F^{-1}(\alpha) \). Similarly, we say that \( q^* \) solves \( \lambda_1 F(q^*) + \lambda_2 F(\lambda_3 q^*) = \alpha \) if \( q^* = \sup\{q : \lambda_1 F(q) + \lambda_2 F(\lambda_3 q) \leq \alpha\} \) (where each \( \lambda_j \geq 0 \) and \( 0 < \alpha < 1 \)). This convention avoids the need to assume that \( \varepsilon \) has a density function.

Henceforth, we let \( K = 1 \) for expository simplicity, although all results are valid if \( K > 1 \). The special case \( K = 0 \) yields a sequence of demands that are independent and identically distributed. The assumption \( K = 1 \) and the notation \( \mu_n = \theta D_{n-1} \) reduces (1.4) to the following demand model used in the remainder of the paper:

\[ D_n = \theta D_{n-1} + \varepsilon_n = \mu_n + \varepsilon_n \]  

(1.5)

We interpret \( \mu_n \) as the deterministic part of demand in period \( n \).
Optimization Objective and Dynamic Program

The following redefinition of the decision variables shortly reduces the dimensionality of the state vector. The idea is to replace the flow variables $z_n$ and $v_n$ with new variables that specify process levels after the period $n$ decisions are implemented.

Let

\begin{align}
y_n &= x_n + z_n \quad \text{(1.6a)} \\
\ s_n &= w_n - p(w_n) - v_n - cz_n - \rho b_n \quad \text{(1.6b)}
\end{align}

The procurement/production quantity, $z_n$, is replaced by the supply level $y_n$ and the dividend, $v_n$, is replaced by $s_n$. The latter is the working capital after paying the dividend, loan interest, and production costs, but before the loan is made and the revenue and inventory costs are realized.

This replacement yields a simple form of the liquidity and logical constraints (1.2) and (1.3):

\begin{align}
b_n + s_n &\geq 0, \quad b_n \geq 0, \quad \text{and} \quad y_n \geq x_n \quad \text{(1.7)}
\end{align}

Let the scalar $\beta$ denote the single-period discount factor ($0 < \beta < 1$). A constant discount factor streamlines the exposition and does not affect the validity of most results (see Section 1.7).

The present value of the dividends net of capital subscriptions is

\begin{equation}
B = \sum_{n=1}^{\infty} \beta^{n-1} v_n,
\end{equation}

Let $H_n$ denote the partial history from period one up to the beginning of period $n$. A *policy* is a non-anticipative decision rule that selects $z_n$, $b_n$, and $v_n$ for each $n$ as
a function of $H_n$. A policy is optimal with respect to the set of initial states $S$ if it maximizes $E(B|H_n)$ for each $H_n$ and $n$ such that $(x_1, \mu_1) \in S$. In the remainder of the paper we characterize an optimal policy.

Substituting $v_n = w_n-p(w_n)-s_n-c(y_n-x_n)-\rho b_n$ (from (1.6b)), $x_{n+1} = (y_n-D_n)^+$ (from (1.3)), and $w_n = s_{n-1}+g(y_{n-1},\mu_{n-1}+\varepsilon_{n-1})$ in (1.8), rearranging terms, and defining

$$L(b,s,y,\mu) = -(1-\beta)s - cy + \beta c E[(y-\mu-\varepsilon)^+]$$
$$+ \beta E\{g(y,\mu+\varepsilon) - p[s + g(y,\mu+\varepsilon)]\} - \rho b$$

yields

$$E(B) = cx_1 + w_1 - p(w_1) + E \sum_{n=1}^{\infty} \beta^{n-1} L(b_n,s_n,y_n,\mu_n)$$

(1.10)

Since a policy maximizes $E(B)$ if and only if it maximizes $E(B') = E(B) - (cx_1+w_1-p(w_1)$, it follows that the maximization of $E(B)$ subject to (1.2) and (1.3) corresponds to the optimization of

$$E(B') = E \sum_{n=1}^{\infty} \beta^{n-1} L(b_n,s_n,y_n,\mu_n)$$

(1.11)

with the following constraints for each $n$:

$$b_n + s_n \geq 0, \quad b_n \geq 0, \quad \text{and} \quad y_n \geq x_n$$

(1.12)

This optimization of the expected present value of dividends corresponds to the following dynamic program:

$$\psi(x,\mu) = \sup_{b,s,y} \{J(b,s,y,\mu) : y \geq x, b \geq 0, b + s \geq 0\}$$

(1.13a)

$$J(b,s,y,\mu) = L(b,s,y,\mu) + \beta E\{\psi[(y-\mu-\varepsilon)^+,\theta(\mu+\varepsilon)]\}$$

(1.13b)
A finite-horizon recursion that corresponds to (1.13) is \( \psi_0(\cdot, \cdot) \equiv 0 \) and

\[
\psi_n(x, \mu) = \sup_{b, s, y} \{ J_n(b, s, y, \mu) : y \geq x, b \geq 0, b + s \geq 0 \}
\]

(1.14a)

\[
J_n(b, s, y, \mu) = L(b, s, y, \mu) + \beta E \{ \psi_{n-1}[(y - \mu - \varepsilon)^+, \theta(\mu + \varepsilon)] \}
\]

(1.14b)

for each \( n = 1, 2, \ldots \). Under reasonable conditions the sequence \( \psi_n \) converges point-wise to \( \psi \) as \( n \to \infty \).

1.4 Preliminary Results

The first result asserts that it is optimal to borrow the smallest amount that satisfies the liquidity constraint (1.7). The proof (omitted) is a straightforward generalization of a proof of a similar result in Li et al. (1997).

**Proposition 1.** In (1.14), \( b = (-s)^+ \) is without loss of optimality.

This result is consistent with the well-known pecking-order hypothesis, which says that firms use internally generated capital before turning to more expensive sources of financing. After internal sources are exhausted, firms will resort to debt.

The following result asserts that the dynamic program value function is a concave function. We omit the inductive proof which starts with \( \psi_0(\cdot, \cdot) \equiv 0 \).

**Proposition 2.** If \( p(\cdot) \) is convex and \( g(\cdot, d) \) is concave (for each possible realization of demand \( d \)), then the dynamic program value function \( \psi_n(\cdot, \cdot) \) in (1.14a) and \( J_n(\cdot, \cdot, \cdot, \cdot) \) in (1.14b) are concave functions on their respective domains \( (n = 1, 2, \ldots) \).

Let \( (b, s, y) = (b_n(\mu), s_n(\mu), y_n(\mu)) \) optimize \( J_n(b, s, y, \mu) \) in (1.14) and (1.14a).
subject only to $b \geq 0$ and $b + s \geq 0$ (i.e., not subject to $y \geq x$). Concavity yields the following result that is a consequence of Propositions 1 and 2 (we omit the proof).

**Proposition 3.** Under the hypotheses of Proposition 2, $(b, s, y) = ((-s_n(\mu))^+, s_n(\mu), \max\{y_n(\mu), x\})$ is optimal in (1.14).

Proposition 3 says that optimal behavior consists of borrowing the smallest amount that satisfies the liquidity constraint and producing nothing if the inventory on hand exceeds the optimal base-stock level.

### 1.5 A Myopic Optimum

Henceforth, we assume that the default penalty function and the net-revenue function are piece-wise linear:

\[
p(x) = (-ax)^+ \quad (a \geq 0) \tag{1.15}
\]

\[
g(y, d) = ry - (r + h)(y - d)^+ \tag{1.16}
\]

Here, $r$ and $h$ are an exogeneous selling price and a holding cost rate. Notice that (1.16) is consistent with the assumption that excess demand is lost. The piece-wise linearity in (1.15) and (1.16) is not necessary for our qualitative conclusions, but it yields simple explicit formulas.

The substitution of (1.15), (1.16), and $b = (-s)^+$ in (1.9) yields

\[
L(b, s, y, \mu) = -(1 - \beta)s - \rho(-s)^+ + (\beta r - c)y - \beta(r + h - c)E[(y - \mu - \varepsilon)^+] - \beta aE\{(r + h)(y - \mu - \varepsilon)^+ - s - ry\}^+ \tag{1.17}
\]
We use the mnemonic \( C \) for “coordinated.” Let \((b, s, y) = (b^C(\mu), s^C(\mu), y^C(\mu))\) maximize

\[
L(b, s, y, \mu) \text{ subject to } b + s \geq 0 \text{ and } b \geq 0.
\]

We focus on a myopic policy \( \pi^C \) that assigns

\[
(b_n, s_n, y_n) = (b^C(\mu_n), s^C(\mu_n), y^C(\mu_n)) \text{ if this triple is feasible, i.e. if } x_n \leq y^C(\mu_n),
\]

and assigns \((b_n, s_n, y_n)\) arbitrarily but feasibly if \(y^C(\mu_n) < x_n (n = 1, 2, \ldots)\). The attraction of \( \pi^C \) is that it would indeed be optimal with respect to \( \{(x, \mu) : x \leq y^C(\mu)\} \) if \( x_1 \leq y^C(\mu_1) \) and \( \pi^C \) induces \( x_n \leq y^C(\mu_n) \) for all \( n \). In accordance with Proposition 1, henceforth we assign \( b^C(\mu) = (-s^C(\mu))^+\).

If the initial inventory were low enough (i.e. \( x_1 \leq y^C(\mu_1) \)), then \((b_1, s_1, y_1) = ((-s^C(\mu_1))^+, s^C(\mu_1), y^C(\mu_1))\) would be feasible in period one. If those decisions were actually executed, then a sufficient condition for \((b_2, s_2, y_2) = ((-s^C(\mu_2))^+, s^C(\mu_2), y^C(\mu_2))\) to be feasible in period two would be that \( \{\mu_n\} \) is non-decreasing over time (with probability one). This is unreasonably restrictive and we do not make this assumption. We focus on \( \pi^C \) because a much weaker assumption implies feasibility, hence optimality, of \((b_n, s_n, y_n) = ((-s^C(\mu_n))^+, s^C(\mu_n), y^C(\mu_n))\) for all \( n = 1, 2, \ldots \).

Let \( y^* \) be the largest value of \( y \) that satisfies

\[
(r + h - c)F(y) + ahF\left(\frac{hy}{r + h}\right) \leq r - c/\beta 
\]

We make the following assumption throughout the remainder of the body of the paper.

\[
F\left(\frac{hy^*}{r + h}\right) < \frac{1 - \beta - \rho}{\beta a} < \frac{\beta r - (1 - \beta - \rho)h - c}{\beta(r + h - c)} 
\]

The rationale for (1.19) begins with Proposition 2 which implies that \((b, s, y) = (b^C(\mu), s^C(\mu), y^C(\mu))\) maximizes a concave function of three variables subject
to two constraints. So one route to a complete solution is the application of the
Karush-Kuhn-Tucker conditions to the optimization problem. This route is followed
in Appendix A.1 and results in a partition of the set of all possible problem parameters
into five subsets. The relevant parameters are $\beta, \rho, c, r, h, a$, and the distribution
function $F(\cdot)$ of the random portion of demand $\varepsilon$. The body of this paper is confined
to the subset that we judge most likely to be encountered in practice, namely (1.19).
The complete solution is given in Appendix A.1.

Condition (1.19) is stronger than necessary for the next result, but it simplifies
the exposition. Following Proposition 5 and in Appendix A.1.3 there are discussions
of the restrictions imposed by (1.19).

An optimal coordinated policy is myopic and stipulates base-stock levels for physical
inventory and retained earnings. Let

$$y^C(\mu) = \mu + f_0 \quad f_0 = F^{-1}\left[\frac{\beta r - (1 - \beta - \rho)h - c}{\beta(r + h - c)}\right] \quad (1.20)$$

$$s^C(\mu) = hf_0 - (r + h)f_1 - r\mu \quad f_1 = F^{-1}\left[\frac{1 - \beta - \rho}{\beta a}\right] \quad (1.21)$$

**Proposition 4.** Assumption (1.19) implies that $\pi^C$ is optimal with respect to $\{(x, \mu) : x \leq y^C(\mu)\}$. That is, if $x_1 \leq y(\mu_1)$ then $(b, s, y) = ((-s^C(\mu_n))^+, s^C(\mu_n), y^C(\mu_n))$ ($n = 1, 2, ...$) is feasible and maximizes (1.11) subject to $x_n \leq y_n, 0 \leq b_n$ and $0 \leq b_n + s_n$
for all $n = 1, 2, ...$.

This result is imbedded in Lemma 1 in Appendix A.1. Although the proof of
Lemma 1 applies here, we sketch the proof of Proposition 4 alone by confirming
feasibility and unconstrained optimality. The latter step concerning $y$ is based on the
following equation imbedded in Lemma 1 whose inverse is the right side of (1.20):

\[
F(y - \mu) = \frac{\beta r - (1 - \beta - \rho)h - c}{\beta(r + h - c)}
\]

Unconstrained optimality concerning \(s\), i.e., (1.21), can be evaluated when \(y_n = \mu_n + f\) for each \(n\), regardless of whether \(f = f_0\) or not. The proof of the following result is contained in the Appendix.

**Proposition 5.** (a) If \(y = \mu + f\) and \(s < 0\), then \(L[(-s)^+, s, \mu + f, \mu]\) is maximized by

\[
s = hf - (r + h)f_1 - r\mu
\]  

(b) If \(y = \mu + f\) and \(s \geq 0\), then \(L[(-s)^+, s, \mu + f, \mu]\) is maximized by

\[
s = hf - (r + h)f_2 - r\mu \quad f_2 = F^{-1}\left[\frac{1 - \beta}{\beta a}\right]
\]  

(c) If \(x_1 \leq \mu_1 + f\) then \(y_n = \mu_n + f\) is feasible for all \(n = 1, 2, ...\)

(d) If \(y_k = \mu_k + f\) and \(s_k = hf - (r + h)f_1 - r\mu_k\) for \(k = n - 1\) and \(k = n\) then

\[
v_n = r(1 - \rho)\mu_n - \rho(r + h)f_1 - a(r + h)[(f - \varepsilon_{n-1})^+ + f_1 - f]^+
+ (r + h\rho)f - (r + h)(f - \varepsilon_{n-1})^+
\]  

In (1.20), \(y^C(\mu)\) equals mean demand plus a fudge factor, \(f_0\), that is constant with respect to firm size but depends on operational, financial and market factors. The operational factors are the holding cost rate \(h\) and the unit production cost \(c\), the financial factors are the discount factor \(\beta\) and the short-term borrowing interest rate \(\rho\), and the market factor is the unit price \(r\) (and \(F\)). In addition to the obvious dependence on the firm size \(\mu\), both \(y^C(\mu)\) and \(s^C(\mu)\) in (1.20) and (1.21) share
the monotonicity properties of the myopic optimal solution in Li et al. (1997) (cf. Proposition 4.2 in Li et al. (1997)). In particular, \( y^C(\mu) \) is nondecreasing with respect to \( r, \beta \) (if \( \rho \leq 1 \)), and \( \rho \), and is nonincreasing with respect to \( c \) (if \( \rho \leq 1 \)) and \( h \) (if \( c/r \leq (1 - \rho)/(2 - \rho - \beta) \)). Similarly, \( s^C(\mu) \) is nondecreasing with respect to \( a, \rho, \) and \( r \), and is nonincreasing with respect to \( c \) and \( h \).

In (1.22) and (1.23) the level of retained earnings diminishes at rate \( r \) as the firm grows; i.e., larger firms borrow more. On the other hand, that level rises at rate \( h \) (the holding cost rate) as the inventory buffer \( f \) grows. The relevance to \( s \) in Proposition 4 is that (1.19) implies that \( s^C(\mu_n) < 0 \) for each \( n \). So (1.22) and \( f = f_0 \) imply \( s = s^C(\mu) \). Later in the paper we discuss the property of (1.24) that the dividend grows with firm size at the rate of \( r(1 - \rho) \).

Proposition 4 is useful because many combinations of parameters \( \beta, \rho, c, r, h, a, \) and \( F(\cdot) \) that are likely to arise in practice satisfy assumption (1.19). In addition to the restrictions on the scalar parameters \( \beta, \rho, c, r, h, \) and \( a \) which are generally satisfied when the coefficients \( r \) and \( a \) are not particularly low and the borrowing interest rate \( \rho \) is not excessively high in comparison to the discount rate \( \beta \), (1.19) restricts the distribution function. For a given \( F(\cdot) \), say normal with a mean of 100 units and a standard deviation of 30 units, (1.19) is satisfied for a wide range of values of \( \beta, \rho, c, r, h, \) and \( a \). The normality assumption is tenable here because the coefficient of variation is less than one-third; so the probability of a negative value of \( \varepsilon \) is very small. We found that the condition is most sensitive to the cost of borrowing \( \rho \) and the default penalty function coefficient \( a \). Values of either one that are high enough to violate (1.19) call for a lower value of \( y \) and a higher value of \( s \) than given in (1.20).
and (1.21).

For example, let \( r = $10 \) and suppose that the firm’s gross margin in 50%; so \( c = \$5 \). Suppose that the annual holding cost rate is \( h = \$1.25 \) (25% of \( c \)), the discount rate \( \beta = 0.8 \), and the borrowing rate \( \rho = 0.80 \). Let \( \varepsilon \) be normally distributed with mean 100 and standard deviation 30. If the default penalty rate is \( a = \$10 \), these values satisfy (1.19), (1.20), and (1.21) with \( y^* = 107.34 \), \( y^C(\mu) = 105.67 + \mu \) and \( s^C(\mu) = -236.43 - 10\mu \). The optimal short-term loan is \( b = 236.43 + 10\mu \). If \( a \in [0.25, 76] \), then (1.19) remains valid, \( y^C(\mu) \) remains the same, and \( s^C(\mu) \) ranges from \(-992.90 - 10\mu \) to \(-10\mu \). If \( h \) were to increase from \( h = \$1.25 \) to \( h = \$4 \) (80% of the production cost \( c \)), then (1.19) would remain valid while \( a \in [0.35, 19.25] \). The optimal supply level would become \( y^C(\mu) = 89.33 + \mu \) and \( s^C(\mu) \) would range from \(-888.89 - 10\mu \) to \(-10\mu \).

### 1.6 Firm Size and the Worth of Coordination

This section examines the influence of firm size on the extent to which it is worthwhile to coordinate production and financial decisions rather than leaving them decentralized. First, we consider the influence of firm size on optimal coordinated decisions. It seems intuitive that the following managed quantities should increase as the firm grows: inventory level, size of short-term loan, and dividend. However, some of the rates of increase may not be intuitive.

In this section we assume that the initial inventory is sufficiently low to permit \( y_1 = \mu_1 + f_0 \), i.e., \( x_1 \leq \mu_1 + f_0 \) and, therefore, that \( y_n = \mu_n + f_0 \) for all \( n \), as stipulated.
by policy \( \pi^C \), is feasible. We refer to \( \pi^C \) as the “optimal coordinated” policy.

The following assertion is an immediate consequence of the roles of \( \mu \) in (1.20) and (1.21) and of \( \mu_k \) in (1.24), and of \( b = (-s)^+ \).

**Proposition 6.** The optimal coordinated policy \( \pi^C \) induces a physical goods base-stock level, short-term loan, and dividend that increase linearly with firm size. The respective growth rates are \( 1, r, \) and \( r(1 - \rho) \). The residual retained earnings decreases with firm size at rate \( r \).

The least intuitive part of this result is that the dividend growth rate is \( r(1 - \rho) \). In the demand model (1.5), \( \mu_n \) is a “sure thing;” that is, the firm’s revenue in period \( n \) is at least \( r\mu_n \) (because \( \varepsilon_n \geq 0 \)). So the “sure” revenue \( r\mu \) grows at rate \( r \) as a function of \( \mu \) (the same rate at which borrowing increases and retained earnings decreases). Policy \( \pi^C \) directs a dividend-maximizing firm to borrow “against” any increase in the sure revenue stream and immediately to pass the net borrowing proceeds to shareholders in the form of a dividend. Since \( \rho \) is the interest rate on short-term loans, the rate of change (with respect to \( \mu \)) of the net borrowing proceeds is \( r(1 - \rho) \).

This phenomenon occurs in practice as asset securitizations and asset-backed borrowing. These actions are frequently utilized by financially sophisticated firms to borrow at favorable rates against future revenue streams.

**Decentralized Operations and Financial Decisions**

The optimal coordinated policy is more profitable than any decentralized policy because the model does not include the infrastructure costs of coordination. In practice,
coordination would be worthwhile only if the difference were to exceed the infrastructure costs. This subsection specifies that difference and relates it to firm size. The preliminary issue is how to assess (1.10), the EPV (expected present value) of the dividends associated with a decentralized policy. The comparable EPV of the optimal coordinated policy is $\psi(x_1, \mu_1)$, namely the value function of the dynamic program at the initial state.

We assign an EPV to decentralized decisions by assuming that production decisions optimize the EPV of net profit, i.e., they are made without considering the consequences for borrowing, dividends, and bankruptcy. Therefore, the production decisions are made under the implicit assumption that the Miller-Modigliani theorem is applicable to the firm. The borrowing and dividend decisions are made subsequently; so financial management takes as given the cash flows induced by the production decisions.

As a result of these behavioral assumptions, operations maximizes the expected value of the following random variable:

$$B^o = \sum_{n=1}^{\infty} \beta^{n-1}[g(y_n, D_n) - cz_n]$$

This is the EPV of the revenues net of inventory- and production-related costs. The substitution of (1.5), (1.16), $z_n = y_n - x_n$, and $x_n = (y_{n-1} - D_{n-1})^+$, and a rearrangement of terms yields

$$B^o = cx_1 + \sum_{n=1}^{\infty} \beta^{n-1}[(r - c)y_n - (r + h - \beta c)(y_n - \mu_n - \varepsilon_n)^+]$$

Since the expected value of the $n^{th}$ summand is a newsvendor objective, when $\mu_n = \mu$
it is maximized by

\[ y = \mu + f_3 \quad f_3 = F^{-1}\left[\frac{r - c}{r + h - \beta c}\right] \] (1.25)

That is, if \( x_1 \leq \mu_1 + f_3 \), then the myopic decisions \( y_n = \mu_n + f_3 \) (for all \( n = 1, 2, ... \)) are feasible and (sub-)optimal. In the decentralized mode, we assume operations makes these decisions.

Next, finance selects short-term loans and dividends to maximize \( E(B) \) given by (1.10) in which \( y_n = \mu_n + f_3 \) and \( b_n = (-s_n)^+ \) (for all \( n = 1, 2, ... \)). The arguments that lead to (1.11) yield the objective of maximizing

\[
E \sum_{n=1}^{\infty} \beta^{n-1} L[(-s_n)^+, s_n, \mu_n + f_3, \mu_n]
\] (1.26)

Proposition 5 implies that the value of \( s \) that maximizes \( L[(-s)^+, s, \mu + f_3, \mu] \) is

\[
s = hf_3 - r\mu - (r + h)f_1
\] (1.27)

if \( s < 0 \). If \( s \geq 0 \), the solution is

\[
s = hf_3 - r\mu - (r + h)f_2
\] (1.28)

Let \( \pi^D \) be a policy that employs (1.25), (1.27), and \( b_n = (-s_n)^+ \) with \( \mu = \mu_n \) if \( x_n \leq \mu_n + f_3 \), and chooses \((b_n, s_n, y_n)\) arbitrarily but feasibly if \( x_n > \mu_n + f_3 \) (the mnemonic D denotes “decentralized”). Let \( y^D(\mu) \) and \( s^D(\mu) \) denote the values of \( y \) and \( s \) in (1.25) and (1.27), respectively, let \( b^D(\mu) = (-s^D(\mu))^+ \), and let \( \Delta \) label the difference between EPVs of the dividends induced by \( \pi^C \) and \( \pi^D \). This is the difference between the maximal values of (1.11) and (1.26). The Appendix contains a proof of the following result.
**Proposition 7.** $\Delta$ does not depend on the initial size of the firm, namely $\mu_1$.

Proposition 7 implies that $\Delta/\mu_1 \to 0$ as $\mu_1 \to \infty$. So it is particularly worthwhile for small firms to coordinate production and financial decisions.

Next we show that the coordinated policy induces lower levels of inventory and retained earnings, and larger short-term loans than the decentralized policy. The Appendix contains the proof of the following result.

**Proposition 8.** (a) $y^C(\mu) \leq y^D(\mu)$, $s^C(\mu) \leq s^D(\mu)$, and $b^C(\mu) \geq b^D(\mu)$.

(b) The rates of change with respect to $\mu$ of the physical goods base-stock level, retained earnings level, short-term loan, and dividend are the same under $\pi^C$ and $\pi^D$.

### 1.7 Financial Distress

This section addresses two issues associated with financial distress. First, we derive the default probability that is induced by the optimal coordinated policy $\pi^C$. Second, we analyze the consequences if bankruptcy forces irrevocable dissolution of the firm; we call this “wipeout bankruptcy.” Thus far in the model, bankruptcy has led to reorganization of the firm with attendant costs; we call this “reorganization bankruptcy.”
Reorganization bankruptcy occurs at the beginning of period \( n + 1 \) if the residual retained earnings is negative, i.e., \( w_{n+1} < 0 \). Since

\[
w_{n+1} = s_n + g(y_n, D_n) = s_n + ry_n - (r + h)(y_n - \mu_n - \varepsilon_n)^+,
\]
default occurs if \( \varepsilon_n \) takes a value in the set \( \{ e : s + ry < (r + h)(y - \mu - e)^+ \} \).

Under the optimal coordinated policy, \( y \) and \( s \) are given by (1.20) and (1.21); so \( y - \mu - e = f_0 - e \) and \( s + ry = (r + h)(f_0 - f_1) > 0 \) with the inequality due to assumption (1.19). Therefore, default occurs if \( \varepsilon_n \) takes a value in

\[
\{ e : s + ry < (r + h)(f_0 - e)^+ \} = \{ e : e < f_0 - (s + ry)/(r + h) \} = \{ e : e < f_1 \}
\]

Let \( q \) denote the probability of default in period \( n \). Since \( q = P\{ e < f_1 \} \) regardless of \( \mu_n \), the default probability is the same every period and is invariant with respect to the size of the firm. So the duration of a sequence of default-free periods does not depend on the size of the firm at the beginning of the sequence and it has a geometric distribution (as in Li et al. (1997)).

**Proposition 9.** Under the optimal coordinated policy, default occurs in a sequence of periods comprising a renewal process. The interval of time between successive defaults has a geometric distribution that does not depend on the firm’s initial size.

**Comments**

(i) This result implies that it is reasonable for the discount factor \( (\beta) \) and the borrowing interest rate \( (\rho) \) to be constants because the default risk is constant.

(ii) The intuition for this result lies in the additive structure of the autoregressive
demand model.

(iii) Li et al. (1997) reach this conclusion for the special case of our demand model with $K = 0$.

(iv) If the reasoning that yields Proposition 9 is employed in the model with wipeout bankruptcy, it yields the conclusion that the lifetime of the firm has a geometric distribution that does not depend on the firm’s initial size.

The Option to Declare Bankruptcy in Wipeout Bankruptcy

Bankruptcy in reality is a complex legal process that we have thus far simplified by charging a default penalty when retained earnings is negative, i.e., $w_{n+1} < 0$ leads to the costs of restructuring debt. In this subsection, we consider the alternative scenario of wipeout bankruptcy. Without the opportunity to restructure debt, the firm has the option to declare bankruptcy either by “looting the treasury” (letting $v_n = w_n$ and $z_n = b_n = 0$) or by “walking away,” i.e., surrendering its assets if the demand in period $n, D_n$, fails to generate sufficient revenue thus causing $w_{n+1} = s_n + g(y_n, D_n) < 0$. The latter is particularly important when the firm relies on external borrowing because one of the model’s implicit assumptions is that the ownership structure of the firm is consistent with a limited liability corporation. That gives the firm the option to return its assets to the lenders for the loans’s notional value at maturity.

Let $q^c(s, y, \mu)$ denote the probability that bankruptcy does not occur in period $n + 1$ if $s_n = s, y_n = y$, and $\mu_n = \mu$. An extension of the argument in Li et al. (1997) is applicable here and shows that dynamic program (1.13) remains valid with wipeout
bankruptcy when (1.13b) is replaced with

\[
J(b, s, y, \mu) = L_w(b, s, y, \mu) + \beta q^c(s, y, \mu)E\{\psi[(y - \mu - \varepsilon)^+, \theta(\mu + \varepsilon)]\}
\]

\[
L_w(b, s, y, \mu) = q^c(s, y, \mu)E[\beta g(y, \mu + \varepsilon) - (1 - \beta)s
\]

\[
- cy + \beta cE[(y - \mu - \varepsilon)^+] - \rho b] - (s + cy + \rho b)(1 - q^c(s, y)),
\]

Observe that \(L_w(b + \mu, s + \mu, y + \mu, \mu)\) is constant with respect to \(\mu\); so \(\Lambda_w(b, s, y) = L_w(b + \mu, s + \mu, y + \mu, \mu)\) is well-defined as is the following dynamic program with \(\Psi(x) = \psi(x - \mu, \mu)\):

\[
\Psi(x) = \sup_{b,s,y}\{M(b,s,y) : y \geq x, b \geq 0, b + s \geq 0\}
\]

\[
M(b,s,y) = \Lambda_w(b,s,y) + \beta q^c(s,y)E(\Psi[(y-\varepsilon)^+])
\]

This dynamic program is formally the same one that is analyzed in (Li et al. (1997), Section 5). Therefore, there are optimal base-stock levels for physical goods inventory and retained earnings, and the lifetime of the firm is geometrically distributed with a parameter that depends only on the optimal base-stock levels (and not on firm size).

The key insights from this section are (a) the default probability is constant with respect to firm size, (b) changes in the bankruptcy mechanism have only a scaling effect on the optimal base-stock levels, and (c) the model with wipeout bankruptcy shares key properties of the model with reorganization bankruptcy and in some ways is easier to analyze.
1.8 Conclusions and Generalizations

We address the coordination of production and financial decisions in a firm that maximizes the expected present value of dividends and that encounters demand that evolves according to an autoregressive process. The firm can borrow funds on a short-term basis, which in turn exposes it to the risk of bankruptcy. We use the deterministic part of demand as the proxy for firm size, $\mu$.

The contributions include the complete characterization of an optimal policy that is myopic and entails base-stock levels for physical goods inventory and retained earnings. The optimal borrowing strategy follows the well-known pecking-order hypothesis. The base-stock level for inventory equals $\mu$ plus a factor, and the one for retained earnings equals $-r\mu$ plus a factor ($r$ is the exogenous price for the firm’s product). Neither factor depends on $\mu$. The optimal policy yields dividends that grow with firm size at the rate of $r(1 - \rho)$ ($\rho$ is the interest rate for borrowing short-term).

Since the infrastructure costs of coordination are not included in the model, it is necessarily better to coordinate production and financial decisions than to decentralize them. We compare the effects of optimal coordinated and decentralized decisions and find repeatedly that the relative improvement of coordination is greater for small firms rather than for large.

It is important to examine the dependence of the conclusions on the assumptions. For example, bankruptcy is a complex financial and legal process, and in most of the paper we assume that the firm’s inability to repay a loan leads to restructuring its debt with associated costs. In Section 1.7 we explain why the replacement of the
restructuring assumption with irrevocable dissolution of the firm would preserve key results.

Following preliminary results (Propositions 1, 2, and 3) we assume particular piece-wise linear expressions for the default penalty and the sales revenue net of inventory-related costs. Most of the qualitative results would be preserved with a general default penalty function that is nonincreasing and convex, and a general net sales revenue function that is concave. However, simple explicit formulas would be lost.

In Section 1.3 we introduce a general autoregressive demand process (1.4), but we employ a first-order process thereafter for expository simplicity (i.e., $K = 1$ after (1.4)). The use of a higher-order process would preserve most of the paper’s qualitative results, but simple explicit formulas would be lost.

In the model the firm pays a constant rate of interest for short-term loans, but the qualitative results would be preserved if, instead, the interest rate were a convex nondecreasing function of the probability of default. However, Section 1.7 shows that this generality is unnecessary because the probability of default remains constant. Similarly, the model uses a constant single-period discount factor, and converting this constant to a function is unnecessary.
Chapter 2

Valuing Adaptive Quality and Shutdown Options in Production Systems

2.1 Introduction

This paper is concerned with the value of operational flexibility and how best to exploit it. We consider two forms of operational flexibility in a production setting: the option to shut down production temporarily in order to avoid losses, and the option to control expected production yield through costly effort. The former is referred to as the temporary shutdown option, the latter as the quality option.

In practice, production yield is influenced by many actions: e.g., maintenance of the production technology, choice of input material and quality control, to name just a few. Some actions have an effect only on the current batch, while other actions
have a longer-term effect, impacting multiple batches. For example, choosing a higher quality raw material for a particular batch only impacts production of that batch. On the other hand, a better maintained production technology will likely result in higher yields in future batches as well. For our purposes, production yield is influenced by effort, which is used as a proxy for a variety of decisions. The cases where effort has short-term and long-term effect are considered separately.

It is clear that revenues, costs or production yield must be uncertain if these options are to play a role. Uncertainty is introduced by supposing that prices and production yield follow continuous time stochastic processes. The manager of the production technology (a firm) is a price-taking profit maximizer.

A real-world example from this class of problems are peaking power plants that generally run only when there is a high demand for electricity. In the United States, this often occurs during hot summer days when the air conditioning load is high. In the unregulated electricity markets, these temporary spikes in demand are accompanied by increases in the price of electricity, making it temporarily profitable to operate these otherwise inefficient plants. The time that a peaker plant operates may be many hours a day or as little as a few hours per year, depending on the condition of the region’s electrical grid. Clearly, in this example, the temporary shutdowns are an important operational component. The efficiency with which these units generate electricity tends to be low and widely varied depending on the factors such as age, maintenance, and the choice of fuel. (Commonly these plants have a dual fuel capability and can run on natural gas or a range of petroleum based products.) Operational decisions that control these factors are the quality option component and their con-
tribution is examined in this paper. Other examples include commodity extraction problems (cf. Brennan and Schwartz (1985) and Enders et al. (2007)). While inspired by real-world applications such as these, the model presented here is intended to be quite general.

When there is no yield uncertainty, then the issue of yield management disappears, and the firm’s only decision in each period is whether or not to exercise the temporary shutdown option. McDonald and Siegel (1985) (hereafter MS) have closely investigated the value of technologies where these options can be exercised in consecutive production periods. In their formulation, the shutdowns and restarts are costless. They show that the value of the technology consists of the sum of independent European exchange options, with successive options giving the holder the right to exchange input for output at a fixed conversion cost at the start of the successive periods. Under assumptions about the dynamics of the input and output prices, they obtain explicit expressions for the value of these exchange options and show that the value of the technology depends critically on the volatilities of input and output prices as well as on their correlation.

When effort is costless, then the production yield is not uncertain, and a special case of our model reduces to the MS model. However, in most production settings, yield enhancements are costly; better maintenance or higher quality input materials are not free. In this case, in each period, the firm has to decide not only on whether to exercise the temporary shutdown option, but if production is to take place, on how much effort to expend. Intuition suggests that as the difference between the output and input prices expands, the firm will expend more effort to enhance the likelihood of
higher yields. Conversely, when the difference between output and input prices drops below some threshold, the profit maximizing firm may choose to expend less effort; or if the gap is sufficiently small, to shut the system down temporarily. The exact amount, of course, depends on the cost of effort; but clearly, it will fluctuate according to the movements in prices of inputs and outputs. Under specific assumptions we are able to characterize an optimal policy completely, and in some cases we are able to establish analytical expressions for the present value of an optimally managed technology.

When a policy with a fixed level of effort in each production period is specified, it is possible to value the technology and to identify the appropriate shutdown and restart operating policy. The value of the production technology operating under the fixed effort policy will be lower than that derived from an adaptive effort policy. The difference between these two values determines the value of the option to control production yield through costly effort – the quality option value. Clearly, this value depends on the output yield under the fixed effort policy. Nonetheless, we show that the value of having an adaptive policy can account for a very large portion of the total value of the production technology. Both of these values increase with price uncertainty. More interestingly, we also investigate the relative contribution of the quality option. We find that this contribution varies with price uncertainty, and, within the model framework described in Section 2.2, we identify sufficient conditions for this contribution to increase or decrease as uncertainty increases. This result provides insight into the value of investing in costly yield management technologies.

Our work is related both to the real options literature and to research on stochastic
maintenance. For an excellent treatment of real options, see the textbook of Dixit and Pindyck (1994), and articles by Brennan and Schwartz (1985), Brennan and Trigeorgis (2000), and Enders et al. (2007). A comprehensive review of real options approaches is provided by Hubbard (1994). Recent research in maintenance is given in Wang (2002) and Wang and Pham (2006). Our contribution is to add yield uncertainty into the paradigm, and to evaluate conditions under which allowing for dynamic adaptive yield management policies adds significant value in the presence of these other managerial options.

The next section outlines a basic model, which is a generalization of the model considered by McDonald and Siegel (1985). The innovation in our model is that production yield is uncertain, but controllable through effort. The impact of the costly effort in the basic model is short-term, an assumption we relax in Section 2.3. This innovation allows us to obtain results that differ in meaningful ways from those obtained by McDonald and Siegel (1985). The section investigates an optimal operating policy and the valuation problem of the entire production technology with the temporary shutdown and quality options and of the quality option alone. It concludes with analysis of the contribution of the quality option in the presence of the temporary shutdown option and how this contribution changes with uncertainty in prices. Section 2.3 extends our results to settings beyond the basic model. The final section concludes.
2.2 Basic Model

The current time is $t_0$. A profit maximizing firm possesses a sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$, where $t_{i+1} - t_i = \Delta t$, $i = 1, 2, \ldots, n-1$, and $t_0 \leq t_1$. The production technology lasts until time $t_{n+1} = t_n + \Delta t$, at which time the salvage value is assumed to be zero. The production time is $\Delta t$.

At each time $t_i$, the firm decides whether to produce between times $t_i$ and $t_{i+1}$. If the decision is to produce, then one unit of raw material is purchased at market price, $I_i$, and present value of the fixed cost of manufacturing for the period, $K$, is incurred.

The production yield is a random variable influenced by the effort level employed at time $t_i$. Let $Q_{i+1}$ represent the quantity of finished goods that can be sold at time $t_{i+1}$ at market price $S_{i+1}$, and let $k' \in [0, 1]$ represent a normalized effort level. The effort levels are ordered so that $k' = 0$ represents no effort leading to the lowest output quantity, none of which can be sold, and $k' = 1$ represents maximum effort, where one unit of output is expected to sell for each unit of input.

The expected value of $Q_{i+1}(k'_i)$, given level of effort $k'_i$ is $k'_i$. The actual value could deviate from this target due to idiosyncratic factors. In particular, we have:

$$Q_{i+1} = H_2(k'_i, z_4(t_{i+1})), \quad (2.29)$$

where $z_4(t_{i+1})$ represents the idiosyncratic production yield factor. The random quantity shocks are assumed to be independent and uncorrelated with all other stochastic processes in our model. For example, $Q_{i+1}$ may follow

$$Q_{i+1} = k'_i e^{-\frac{1}{2}\sigma_Q^2 + \sigma_Q z_4(t_{i+1})}, \quad (2.30)$$
where $z_4(t_{i+1})$ is a standard normal random variable.

To attain the level of effort $k'_i$, a cost $c(k'_i)$ is incurred. We assume that the function, $c(k'_i)$, is convex and non-decreasing in $k'_i$. The cost is assessed at the beginning of each production period.

Both input and output commodities are traded in complete and arbitrage-free markets. Their prices follow geometric Brownian motion with:

\[
\begin{align*}
\frac{dI}{I} &= (\alpha_I - q_I)dt + \sigma_I dz_I, \\
\frac{dS}{S} &= (\alpha_S - q_S)dt + \sigma_S dz_S,
\end{align*}
\]

where $\alpha_I$ and $\alpha_S$ are the expected rates of return of the input and output, $q_I$ and $q_S$ are convenience yields, $\sigma_I$ and $\sigma_S$ are volatilities. Let $\rho$ denote the instantaneous correlation between input and output prices. The stochastic processes (2.29) and (2.31) are modeled on a complete probability space $(\Omega, \mathcal{F}, P)$ equipped with filtration $\{\mathcal{F}_t\}, t \in [t_0, t_{n+1}].$

We also assume that both the raw material purchases and changes in effort require no preparatory lead times and that the firm is atomistic in that its actions cannot influence the raw material or the finished goods market prices. Finally, the firm operates in an economy with a constant risk free rate $r$.

### 2.2.1 Analysis

Under the basic model assumptions, the firm’s time $t_i$ expected profit is:

\[
\pi(I_i, S_i; k'_i, \delta'_i) = E^P_0[(Q_{i+1}S_{i+1}M(t_i, t_{i+1})\delta'_i) - [(I_i + c(k'_i) + K)]\delta'_i],
\]

(2.32)
where \( M(t_i, t_{i+1}) \) is the pricing kernel or stochastic discount factor, and \( \delta'_i = \{0, 1\} \) denotes the firm’s time \( t_i \) production decision. We use a superscript on the expectation to emphasize the probability measure with respect to which the expectation is computed and a subscript to indicate filtration.

By independence of \( z_4(t_{i+1}) \) with other factors, the expectation above can be written as:

\[
E^P_{t_i}[Q_{i+1}S_{i+1}M(t_i, t_{i+1})\delta'_i] = k'_i E^P_{t_i}[S_{i+1}M(t_i, t_{i+1})\delta'_i].
\]

Since \( I_i \) and \( S_i \) for all \( i = 1, 2, \ldots, n \) are prices of traded goods, to avoid riskless arbitrage, there must exist an equivalent martingale measure, \( Q \), under which:

\[
\begin{align*}
\frac{dI}{I} &= (r - q_I)dt + \sigma_I dw_I \quad (2.33a) \\
\frac{dS}{S} &= (r - q_S)dt + \sigma_S dw_S \quad (2.33b)
\end{align*}
\]

with \( E^Q_{t_i}[dw_IDw_S] = \rho dt \), and:

\[
E^P_{t_i}[S_{i+1}M(t_i, t_{i+1})] = e^{-r\Delta t} E^Q_{t_i}[S_{i+1}] = e^{-q_S\Delta t} S_i, \quad (2.34)
\]

where, in (2.33), \( dw_I \) and \( dw_S \) are \( Q \) measure Brownian motions.

Combining (2.32) and (2.34) and letting

\[
g(k'_i) = c(k'_i) + K, \quad (2.35)
\]

the time \( t_i \) expected profit then can be written as:

\[
\pi(I_i, S_i; k'_i, \delta'_i) = \delta'_i [k'_i e^{-q_S\Delta t} S_i - I_i - g(k'_i)]. \quad (2.36)
\]

We continue our analysis with the derivation of an optimal policy, the temporary shutdown and effort decisions that maximize the expected value the firm’s profit.
from production. Using this policy, we then consider the valuation problem of this production technology and conclude with the main result of this section, an analysis of the quality option contribution, which examines the role of the quality option in the presence of the temporary shutdown option.

2.2.2 An Optimal Policy

Since the expected production yield is only influenced by the time $t_i$ effort level $k'_i$, the optimal policy is myopic. Let $k'^*_i$ be the level of $k'_i$ at which the right-hand side of (2.36) attains maximum with $\delta'_i = 1$. That is

$$k'^*_i(S_i) = \arg \max_{k'^i \in [0,1]} \{ k'^i e^{-qS\Delta t S_i} - I_i - g(k'^i) \}. \tag{2.37}$$

A profit-maximizing firm opts to produce whenever there exists $k'^*$ given by (2.37) such that:

$$e^{-qS\Delta t k'^* S_i} - I_i - g(k'^*) \geq 0. \tag{2.38}$$

Therefore, the time $t_i$ expected profit from an optimally managed technology can be written as

$$\pi(I_i, S_i) = \max\{ k'^* e^{-qS\Delta t S_i} - I_i - g(k'^*), 0\}. \tag{2.39}$$

For the remainder of this section, we assume that the cost function, $g(k')$, is of the following form:

$$g(k') = K + ck'^2, \quad c \geq 0 \tag{2.40}$$

where $K$ is the fixed cost of manufacturing for the period and $c$ is the cost of effort
coefficient. This function is not necessary for our qualitative conclusions, but it yields simple explicit formulas.

Substituting (2.40) for $g(k')$ in (2.37) yields

$$k'^*(S_i) = \arg \sup_{k' \in [0, 1]} \left\{ k'e^{-qs\Delta t}S_i - I_i - K - ck'^2 \right\} = \min \left\{ \frac{S_i e^{-qs\Delta t}}{2c}, 1 \right\}, \quad (2.41)$$

where the final result on the right-hand side of (2.41) follows from the first-order condition with respect to $k'$; second-order condition with respect to $k'$ yields concavity.

With (2.41), the time $t_i$ expected profit given by (2.39) then becomes

$$\pi(I_i, S_i) = \begin{cases} 
\max \left\{ \frac{e^{-2qs\Delta t}S_i^2}{4c} - I_i - K, 0 \right\} & \text{if } S_i < 2ce^{qs\Delta t}; \\
\max \left\{ e^{-qs\Delta t}S_i - I_i - K - c, 0 \right\} & \text{if } S_i \geq 2ce^{qs\Delta t}.
\end{cases} \quad (2.42)$$

We now present the optimal operating policy that corresponds to (2.40).

**Proposition 10.** Assume $g(k')$ is given by (2.40). Then the optimal exercise policy for the temporary shutdown and quality options is given as follows:

(i) $I_i < c - K$

$$\delta'^* = 0 \quad \text{if } S_i < S_1^*; \quad (2.43a)$$

$$\delta'^* = 1, \quad k'^* = \frac{e^{-qs\Delta t}S_i}{2c} \quad \text{if } S_1^* \leq S_i < S_2^*; \quad (2.43b)$$

$$\delta'^* = 1, \quad k'^* = 1 \quad \text{if } S_2^* \leq S_i; \quad (2.43c)$$

(ii) $I_i \geq c - K$

$$\delta'^* = 0 \quad \text{if } S_i < S_3^*; \quad (2.44a)$$

$$\delta'^* = 1, \quad k'^* = 1 \quad \text{if } S_3^* \leq S_i; \quad (2.44b)$$
where \( S_1^* = 2e^{qs\Delta t}\sqrt{c(I_i + K)} \), \( S_2^* = 2ce^{qs\Delta t} \), \( S_3^* = e^{qs\Delta t}(c + I_i + K) \).

**Proof.** See Appendix.

First, consider the case when input prices are below the threshold \( c - K \). The temporary shutdown option will be exercised if \( S_i < 2e^{qs\Delta t}\sqrt{c(K + I_i)} \). Notice that as the convenience yield, \( q_s \), cost of effort coefficient, \( c \), or the fixed conversion cost, \( K \), increase, the likelihood of exercising the option increases. For \( K = 0 \), the condition reduces to \( S_i^2/I_i < 4ce^{2qs\Delta t} \). If prices are lognormal, then, viewed from time \( t_0 \), the random variable \( S_i^2/I_i \) is lognormal and its variance decreases as the correlation between the price of output \( S_i \) and the cost of input \( I_i \) increases. Hence the temporary shutdown option component increases in value as the price correlation decreases. If \( S_1^* < S_i < S_2^* \), then the temporary shutdown option expires worthless, and the quality option is exercised with optimal effort, \( k'^* \), at an intermediate value. In particular, \( 0 < k'^* = (e^{-qs\Delta t}S_i)/2c < 1 \). Notice that when the quality option is exercised, the effort level increases with the output price. The lowest effort occurs when \( S_i = S_1^* \), where the firm is indifferent between exercising the temporary shutdown option or exercising the quality option with \( k'^* = \sqrt{(I_i + K)/c} \). Finally, if the price of output \( S_i \) exceeds \( S_2^* \), then the quality option should be exercised at its highest value, \( k'^* = 1 \).

Second, consider the case where \( I_i \geq c - K \). There is no feasible point where the quality option is exercised with \( 0 < k'^* < 1 \). In this case, either the temporary shutdown option is exercised or the quality option is exercised with \( k'^* = 1 \).
2.2.3 Valuation

Since an optimal policy in the basic model is myopic, then the value of the entire sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$ viewed from time $t_0$, say $V_0(I_0, S_0)$, is:

$$V_0(I_0, S_0) = \sum_{i=1}^{n} v_0^{(i)}(I_0, S_0), \text{ where}$$

$$v_0^{(i)}(I_0, S_0) = e^{-r(t_i - t_0)} Q_t[\pi(I_i, S_i)]. \text{ (2.45b)}$$

With the optimal policy given in Proposition 10 and the cost of input $I$ constant, the expression for $v^{(i)}(S_0)$ has a Black and Scholes (1973) format. For the simpler case where $I \geq c - K$, $v^{(i)}(S_0)$ is given by the European call option formula of Merton (1973). For the case where $I < c - K$, the analytical expression for $v^{(i)}(S_0)$ is slightly more complicated, but it still consists of Black-Scholes terms. This result is presented next.

When input prices are stochastic, analytical approximations or numerical procedures, as in Carmona and Durrleman (2003) or Li et al. (2008), for example, are necessary to obtain solutions. We use these approximations later in our numerical experiments.

**Proposition 11.** Assume the output price, $S$, follows (2.33b) and the cost of input, $I$, is constant. Then the market value of the sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$ viewed from time $t_0$, when the effort can adapt to market conditions is

$$V_0(S_0) = \sum_{i=1}^{n} v_0^{(i)}(S_0),$$
where $X = I + K$ and

(i) for $X < c$

$$\nu_0^{(i)}(S_0) = \frac{S_0^2}{4c} e^{(r - 2qs + \sigma^2/2)(t_i - t_0) - 2qs \Delta t} \left[ N(d_{2l}^{(i)}) - N(d_{2u}^{(i)}) \right] - e^{-r(t_i - t_0)} X N(d_{3l}^{(i)})$$

$$+ e^{-qS(t_i - t_0 + \Delta t)} S_0 N(d_{1u}^{(i)}) - e^{-r(t_i - t_0)} c N(d_{3u}^{(i)}); \quad (2.46)$$

$$d_{1l}^{(i)} = \log \left[ e^{-qs\Delta t} S_0/2\sqrt{cX} \right] + \frac{(r - qs + \sigma^2/2)(t_i - t_0)}{\sigma S \sqrt{t_i - t_0}}; \quad d_{2l}^{(i)} = d_{1l}^{(i)} + \sigma S \sqrt{t_i - t_0};$$

$$d_{1u}^{(i)} = \log \left[ e^{-qs\Delta t} S_0/2c \right] + \frac{(r - qs + \sigma^2/2)(t_i - t_0)}{\sigma S \sqrt{t_i - t_0}}; \quad d_{3l}^{(i)} = d_{1l}^{(i)} - \sigma S \sqrt{t_i - t_0};$$

$$d_{2u}^{(i)} = d_{1u}^{(i)} + \sigma S \sqrt{t_i - t_0}; \quad d_{3u}^{(i)} = d_{1u}^{(i)} - \sigma S \sqrt{t_i - t_0};$$

(ii) for $X \geq c$

$$\nu_0^{(i)}(S_0) = e^{-qS(t_i - t_0 + \Delta t)} S_0 N(d_{4l}^{(i)}) - e^{-r(t_i - t_0)} (c + X) N(d_{5l}^{(i)}); \quad (2.47)$$

$$d_{4l}^{(i)} = \log \left[ e^{-qs\Delta t} S_0/(c + X) \right] + \frac{(r - qs + \sigma^2/2)(t_i - t_0)}{\sigma S \sqrt{t_i - t_0}};$$

$$d_{5l}^{(i)} = d_{4l}^{(i)} - \sigma S \sqrt{t_i - t_0}.$$

Proof. See Appendix. \qed

In Proposition 11, when $I \geq c - K$ and $\Delta t \to 0$, the model reduces to the simple temporary shutdown option considered by McDonald and Siegel (1985). In this case, the cost of effort coefficient $c$ is so small that intermediate values for effort are never optimal and the shutdown option alone acts as the yield management tool. Therefore the quality option has no value.
In contrast, when $I < c - K$, the cost of improving quality is high, and for this case the firm finds value in having the flexibility to adapt its costly effort dynamically in response to changing market conditions.

From the right side of (2.41), an optimal effort level $k^*$ is increasing in $S_i$. Differentiation of (2.46) and (2.47) with respect to $S_0$ yields

$$\frac{\partial V_0(S_0)}{\partial S_0} = \sum_{i=1}^{n} \frac{S_0}{2c_e} e^{(r-2q_S+S_S^2)(t_i-t_0)-2q_S\Delta t} [N(d_{2i}) - N(d_{2u})] + e^{-q_S(t_i-t_0+\Delta t)} N(d_{4i})$$

if $X < c$;

$$\sum_{i=1}^{n} e^{-q_S(t_i-t_0+\Delta t)} N(d_{4i})$$

if $X \geq c$.

It is straightforward to verify that both of these expressions are non-negative, so that the production technology value is non-decreasing in output prices. This property, together with the fact that higher output prices induce higher effort levels, holds under more general conditions. This is explored in Section 2.3.2.

Figure 2.1 illustrates the behavior of the market value of the technology, $V_0(S_0)$, as the cost of effort coefficient, $c$, uncertainty, $\sigma_S$, and the number of periods, $n$. As expected, the value increases as the cost of effort coefficient, $c$, decreases, as the volatility, $\sigma_S$, increases, and as the number of periods, $n$, increases.

When the cost of input is stochastic, numerical procedures must be used to compute the expectation in equation (2.45). For this case, Figure 2.2 shows the behavior of the market value of the technology as the number of periods, $n$, increase, first for different values of $c$, and then for different values of correlation between the market prices of input and output, $\rho$.

As can be seen, stochastic input causes the value of the technology to drop relative
to the deterministic case when correlation is positive. For the value of $\sigma_I = 0$, the exhibit in Panel (a) reduces to Panel (a) of Figure 2.1.

Finally, Panels (c) and (d) of Figure 2.2 show how the value of the production technology changes for the special case when there are twenty production periods remaining and the volatilities $\sigma_S$ and $\sigma_I$ vary at different levels of correlation $\rho$. Specifically, value increases with volatility and decreases with correlation $\rho$.

Figure 2.1: The market value of the sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$ viewed from time $t_0$, $V_0(S_0)$, when the effort can adapt to market conditions, the cost of input $I$ is constant and $K + I < c$ with parameters $S_0 = $50, $q_S = 0.02$, $I = $25, $K = $5, $c = $40, $r = 0.05$, $\sigma_S = 0.45$, $t_1 - t_0 = 0.25$ years, and $\Delta t = 0.25$ years.

2.2.4 Quality Option Contribution

McDonald and Siegel (1985) demonstrate that the ability to shut down and restart production in a single mode adds significantly to the value of the entire firm. Indeed, without considering the value of the temporary shutdown option, peaking power
Figure 2.2: The market value of the sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$ viewed from time $t_0$, $V_0(S_0, I_0)$, when the effort can adapt to market conditions, the cost of input $I_i$ is stochastic and $X < c$ with parameters $S_0 = $50, $q_S = 0.02$, $I_0 = $25, $q_I = 0.01$, $K = $5, $c = $40, $r = 0.05$, $\sigma_S = 0.45$, $\sigma_I = 0.35$, $\rho = 0.2$, $t_1 - t_0 = 0.25$ years, and $\Delta t = 0.25$ years.

plants mentioned in the introduction as a motivating example, would not likely be ever built. This is because the traditional NPV value for these units would typically be negative.

But how important is the quality option in the presence of the temporary shutdown option? This is useful to understand when considering a costly yield management upgrade to an existing single mode production system.

Consider a policy that has a temporary shutdown component and only one mode
of operation. Specifically, if production takes place, the level of effort is fixed at $k'$. In this special case, it is easy to show that the value of the sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$ viewed from time $t_0$ is:

$$V_{fix}(S_0|k') = \sum_{i=1}^{n} v_{fix}^{(i)}(S_0|k');$$

$$v_{fix}^{(i)}(S_0|k') = e^{-qS(t_i-t_0+\Delta t)k'S_0}N(d_1^{(i)}) - e^{-r(t_i-t_0)(ck'^2 + X)}N(d_2^{(i)}); \quad (2.48)$$

$$d_1^{(i)} = \log \left[ k' e^{-qS\Delta t}S_0/(ck'^2 + X) \right] + \frac{(r - qS + \sigma^2/2)(t_i - t_0)}{\sigma S \sqrt{t_i - t_0}};$$

$$d_2^{(i)} = d_1^{(i)} - \sigma S \sqrt{t_i - t_0};$$

and where $X = I + K$.

We define the value of operational flexibility, or equivalently, the value of the quality option, $V_{flex}(S_0|k')$, say, to be the difference between $V_0(S_0)$ and $V_{fix}(S_0|k')$. Clearly, $V_{flex}(S_0|k') \geq 0$.

**Proposition 12.** Assume the output price, $S$, follows (2.33b) and the cost of input, $I$, is constant. Then the market value of the quality option at times $t_1 < t_2 < \cdots < t_n$ viewed from time $t_0$ is:

$$V_{flex}(S_0|k') = \begin{cases} 
\sum_{i=1}^{n} v_{flex}^{(i)}(S_0, k') & \text{if } I < c - K; \\
0 & \text{if } I \geq c - K; 
\end{cases}$$

where

$$v_{flex}^{(i)}(S_0|k') = v_{0}^{(i)}(S_0) - v_{fix}^{(i)}(S_0|k') \quad (2.49)$$

and where $v_{0}^{(i)}(S_0)$ and $v_{fix}^{(i)}(S_0|k')$ are given by (2.46) and (2.48) respectively.

**Proof.** Omitted. \[\square\]
It is straightforward to show that as the volatility of the market price of output, \( \sigma_S \), increases, both \( v_0^{(i)}(S_0; \sigma_S) \) and \( v_{flex}^{(i)}(S_0; \sigma_S | k_f') \) increase in absolute value. This agrees with our standard option pricing theory intuition. Now, let

\[
R^{(i)}(S_0; \sigma_S | k_f') = \frac{v_{flex}^{(i)}(S_0; \sigma_S | k_f')}{v_0^{(i)}(S_0; \sigma_S)}
\]

denote the relative value of the quality option with respect to the full market value of the option to produce at time \( t_i \). It can be thought of as measuring the price of \( v_{flex}^{(i)}(S_0; \sigma_S | k_f') \) in the units of \( v_0^{(i)}(S_0; \sigma_S) \). It follows that \( 0 \leq R^{(i)}(\cdot) \leq 1 \). A value of \( R^{(i)}(\cdot) \) close to one indicates that nearly the entire value of the option to produce can be attributed to the quality option. A value \( R^{(i)}(\cdot) \) close to zero indicates that the temporary shutdown option is the most valuable component of \( v_0^{(i)}(S_0; \sigma_S) \).

Panel (a) of Figure 2.3 illustrates that depending on the cost of effort coefficient, \( c \), the relative value of the quality option can be substantial, ranging from 45% to 98% of the entire value of the option to produce for \( k_f' \leq 0.6 \) and \( c = 40 \). This highlights just how much value the quality option may add to the overall option to produce in the presence of the temporary shutdown option.

One might expect that as output price volatility increases, the relative value of the quality option changes. For example, if \( R^{(i)}(S_0; \sigma_S | k_f') \) decreases in \( \sigma_S \), then

\[
R^{(i)}(S_0; \sigma_S + \varepsilon | k_f') < R^{(i)}(S_0; \sigma_S | k_f'), \quad \varepsilon > 0.
\]

Using (2.49), this can be written as

\[
\frac{v_0^{(i)}(S_0; \sigma_S + \varepsilon) - v_{flex}^{(i)}(S_0; \sigma_S + \varepsilon | k_f')}{v_0^{(i)}(S_0; \sigma_S + \varepsilon)} < \frac{v_0^{(i)}(S_0, \sigma_S) - v_{flex}^{(i)}(S_0; \sigma_S | k_f')}{v_0^{(i)}(S_0; \sigma_S)}, \quad \varepsilon > 0,
\]
which, for small values of $\varepsilon > 0$, can be written as

$$\frac{v_0^{(i)}(S_0; \sigma_S) + \varepsilon \phi_0^{(i)}(S_0; \sigma_S) - v_{fix}^{(i)}(S_0; \sigma_S|k_f') - \varepsilon \phi_{fix}^{(i)}(S_0; \sigma_S|k_f')}{v_0^{(i)}(S_0; \sigma_S) + \phi_0^{(i)}(S_0; \sigma_S)} < \frac{v_0^{(i)}(S_0; \sigma_S) - v_{fix}^{(i)}(S_0; \sigma_S|k_f')}{v_0^{(i)}(S_0; \sigma_S)}.$$ 

Assuming $v_{fix}^{(i)}(S_0; \sigma_S|k_f') > 0$, $v_0^{(i)}(S_0; \sigma_S) > 0$, the above inequality is equivalent to

$$\frac{\phi_{fix}^{(i)}(S_0; \sigma_S|k_f')}{v_{fix}^{(i)}(S_0; \sigma_S|k_f')} > \frac{\phi_0^{(i)}(S_0; \sigma_S)}{v_0^{(i)}(S_0; \sigma_S)} \quad (2.50)$$

and where $\phi_{fix}^{(i)}(S_0|k_f') > 0$ and $\phi_0^{(i)}(S_0) > 0$ are the first derivatives of $v_{fix}^{(i)}(S_0|k_f')$ and $v_0^{(i)}(S_0)$ with respect to $\sigma_S$ and are typically referred to as the vegas. These are given by:

$$\phi_0^{(i)}(S_0) = \frac{S_0^2}{2c} e^{(r-2q_S+\sigma_S^2)(t_1-t_0)-2q_S\Delta t(t_i-t_0)} \sigma_S \left[ N(d_{2i}^{(i)}) - N(d_{2u}^{(i)}) \right]$$ 

$$+ \frac{S_0^2}{2c} e^{(r-2q_S+\sigma_S^2)(t_1-t_0)-2q_S\Delta t[n(d_{2i}^{(i)}) - n(d_{2u}^{(i)})]} \sqrt{t_i-t_0} + e^{-q_S(t_i-t_0+\Delta t)} S_0 n(d_{1u}^{(i)}) \sqrt{t_i-t_0};$$

$$\phi_{fix}^{(i)}(S_0|k_f') = k_f' S_0 e^{-q_S(t_i-t_0+\Delta t)} n(d_1^{(i)}) \sqrt{t_i-t_0}; \quad (2.52)$$
where \( n(\cdot) \) denotes the probability density function of the standard normal distribution.

Note the right side of (2.50) does not depend on \( k'_f \). We use this observation to characterize the behavior of \( R^{(i)}(S_0; \sigma_S|k'_f) \) as a function of \( \sigma_S \) and of \( k'_f \).

**Proposition 13.** Assume the cost of input, \( I \), is constant, \( I < c - K \) and the output price, \( S \), follows (2.33b). Then there exist \( 0 < k_L^{(i)} < k_U^{(i)} < 1 \) such that

\[
R^{(i)}(S_0; \sigma_S + \varepsilon|k'_f) < R^{(i)}(S_0; \sigma_S|k'_f), \quad \varepsilon > 0 \quad \text{if} \quad k'_f \in (0, k_L^{(i)}) \cup (k_U^{(i)}, 1]; \tag{2.53a}
\]

\[
R^{(i)}(S_0; \sigma_S + \varepsilon|k'_f) \geq R^{(i)}(S_0; \sigma_S|k'_f), \quad \varepsilon > 0 \quad \text{if} \quad k'_f \in [k_L^{(i)}, k_U^{(i)}]; \tag{2.53b}
\]

where \( k_L^{(i)} \) and \( k_U^{(i)} \) are the two values of \( k'_f \) such that

\[
\frac{\phi^{(i)}_{fix}(S_0; \sigma_S|k'_f)}{v^{(i)}_{fix}(S_0; \sigma_S|k'_f)} = \frac{\phi^{(i)}_{0}(S_0; \sigma_S)}{v^{(i)}_{0}(S_0; \sigma_S)}. \tag{2.54}
\]

**Proof.** See Appendix. \( \square \)

To illustrate the result in Proposition 13, reconsider our earlier example with \( n = 20 \) potential production periods. By solving (2.54) for \( i = 1, 2, \ldots, 20 \), we obtain results shown in Figure 2.4, which clearly indicate the numbers \( k_L^{(i)} \) and \( k_U^{(i)} \) and how these change as a function of \( i \). For \( i = 1 \), for example, the relative value of the quality option increases with volatility \( \sigma_S \) for all values of \( k'_f \) between 0.44 and 0.68 and decreases otherwise. The actual changes in the quality option relative value for \( k'_f = 0.35, 0.5 \) and 0.98 are illustrated on Panel (b) of Figure 2.3. The example assumes that the firm invests in the quality option at \( \sigma_S = 0.45 \) and then it experiences a change in its relative contribution as \( \sigma_S \) increases. In particular, for the values of \( \sigma_S = 0.45 \) and \( k'_f = 0.98 \), the firm finds it extremely valuable to be able to
adjust effort dynamically in response to the changes in the market prices of output. However, as $\sigma_S$ doubles, the relative contribution of the quality option is nearly cut in half. That is, the simple “on/off” option as considered in McDonald and Siegel (1985) becomes the relatively more valuable yield management tool.

Figure 2.4: Bounds $k_L^{(i)}$ and $k_U^{(i)}$ defined in Proposition 13 with parameters $S_0 = \$50$, $q_s = 0.02$, $I = \$25$, $K = \$5$, $c = \$100$, $r = 0.05$, $\sigma_S = 0.45$, $t_1 - t_0 = 0.25$ years, $\Delta t = 0.25$ years, $i = 1, 2, \ldots, n$, and $n = 20$.

Proposition 13 also yields an analogous result for the $n$-period problem. Define

$$R(S_0; \sigma_S|k_f') = \frac{V_{flex}(S_0; \sigma_S|k_f')}{V_0(S_0; \sigma_S)},$$

where $V_0(S_0)$ and $V_{flex}(S_0|k_f')$ are defined in Propositions 11 and 12 respectively.

**Proposition 14.** Assume the cost of input, $I$, is constant, $I < c - K$ and the output price, $S$, follows (2.33b). Then on the interval $(0, 1)$, there exist $k_L$, $k_L^*$, $k_U^*$, and $k_U$. 

48
such that

\[ R(S_0; \sigma_S + \varepsilon|k'_f) < R(S_0; \sigma_S|k'_f), \quad \varepsilon > 0 \quad \text{if} \quad k'_f \in (0, k_L) \cup (k_U, 1); \]

\[ R(S_0; \sigma_S + \varepsilon|k'_f) \geq R(S_0; \sigma_S|k'_f), \quad \varepsilon > 0 \quad \text{if} \quad k'_f \in [k_L, k_U]; \]

where \( k_L = \min\{k_L^{(i)} : i = 1, 2, \ldots, n\}; \) \( k_L = \max\{k_L^{(i)} : i = 1, 2, \ldots, n\}; \) \( k_U = \min\{k_U^{(i)} : i = 1, 2, \ldots, n\}; \) and \( k_U = \max\{k_U^{(i)} : i = 1, 2, \ldots, n\}. \)

Proof. See Appendix. \( \square \)

The application of Proposition 14 is readily seen in Figure 2.4 – for \( n = 20, k_L = 0.32, k_L = 0.44, k_U = 0.68, \) and \( k_U = 0.94. \) Therefore the relative value of the quality option increases with volatility \( \sigma_S \) for all values of \( k'_f \) between 0.44 and 0.68 and decreases for all values of \( k'_f \) smaller than 0.32 and greater than 0.94. The directional change in the relative value of flexibility for the remaining values of \( k'_f \) is not as analytically tractable. However, it is easily checked numerically.

In summary, for this model we have seen that, through effort, the firm can only improve its expected yield for the period. Here, an optimal policy is myopic and the market value of the option to produce can frequently be given in a closed form.

The key results include the absolute and relative values of the quality option in the presence of the temporary shutdown option. The quality option is essentially worthless when the cost of effort coefficient, \( c, \) is low; its absolute value increases with output price uncertainty, \( \sigma_S; \) and, finally, its relative value \( R(S_0; \sigma_S|k'_f) \) may increase or decrease with \( \sigma_S \) – this depends on the mode, \( k'_f, \) the system operates in without the quality option.
2.3 Model Extensions

We now explore the role of the quality and the temporary shutdown options in a setting where costly effort may influence production yield beyond a single period. An important example of this type of costly effort is mechanical or preventive maintenance. It is reasonable to expect that a production process that undergoes substantial maintenance at time $t_i$ is not only more likely to produce higher quantity of finished goods at time $t_{i+1}$, but also at subsequent times. In contrast, in the earlier model, effort did not impact production yield beyond the current time period.

To capture the long-term effect of the costly effort, we introduce a new state variable $k_i \in [0, 1]$, which represents a normalized state in which the production system enters time $t_i$. That is, $k_i$ is now the state of the system at time $t_i$ before the firm exerts the costly effort $k_i'$. The states $k_i$ are ordered so that $k_i = 1$ represents the best state and $k_i = 0$ the worst. To attain the level of effort $k_i'$, a cost $c(k_i, k_i') + b(k_i')$ is now assessed at time $t_i$. The function $c(k_i, k_i')$ is convex, non-increasing in $k_i$, and non-decreasing in $k_i'$. The function, $b(k_i')$, is convex and non-increasing in $k_i'$.

At time $t_i$, a decision to produce with the effort level $k_i'$ does not imply that the system will end the period in that same state. That only occurs if the firm decides not to produce at time $t_i$, in which case the system simply remains in a warm standby in the same state. The firm is assumed to exert no maintenance effort during periods of no production. In particular, the state of the system at time $t_{i+1}$, $k_{i+1}$, is:

$$k_{i+1} = \begin{cases} H_1(k_i', z_3(t_{i+1})) & \text{if } \delta_i' = 1; \\ k_i & \text{if } \delta_i' = 0; \end{cases} \quad (2.55)$$
where $z_3(t_{i+1})$ is some random variable. The random shocks to the state over successive time periods are assumed to be independent, idiosyncratic, and uncorrelated with all other stochastic processes in our model. Moreover, we assume that the function $H_1(\cdot, \cdot)$ satisfies

\[
\{k_{i+1} | k_{i+1}^\prime \} \succ_{st} \{k_{i+1} | k_i \}, \quad \forall \ k_{i+1}^\prime > k_i.
\]  

(2.56)

From equation (2.56), the state of the system at time $t_{i+1}$, $k_{i+1}$, is stochastically increasing in effort, $k_i^\prime$, so that a better maintained process degrades less and the firm needs to exert less effort to attain the same expected yield in the next production period.

Some production processes also require that the firm incur a fixed cost, say $c_{on}$, if the decision to produce at time $t_i$ follows a temporary shutdown decision at time $t_{i-1}$. Therefore, a variable $\delta_i = \{0, 1\}$ indicating whether the process is “on” or “off” at time $t_i$ is also included in the state space. A temporary production shutdown to warm standby is assumed costless. The remaining assumptions and notation are the same as in Section 2.2.

### 2.3.1 Analysis

To streamline notation, let the state variables be $s_i = (I_i, S_i, k_i, \delta_i)$, and the decision variables, representing the temporary shutdown and quality options be $a_i = (\delta_i^\prime, k_i^\prime)$. As before, letting

\[
g(k_i, k_i^\prime) = c(k_i, k_i^\prime) + b(k_i^\prime) + K,
\]
the time $t_i$ expected profit now becomes:

$$\pi(s_i; a_i) = \delta_i'[k_i'e^{-qs\Delta t}S_i - I_i - g(k_i', k'_i)] - c_{on}(-\delta_i' - \delta_i)^+, \quad (2.57)$$

and an optimal policy, $a^*_1, a^*_2, \ldots, a^*_n$, satisfies the following optimality equation:

$$J_i(s_i) = \pi(s_i; a_i^*) + e^{-r\Delta t}E^Q_t[J_{i+1}(I_{i+1}, S_{i+1}, k_{i+1}, \delta^*_i)|a_i^*]] \quad (2.58)$$

$$= \sup_{0 \leq k' \leq 1, \delta' = 0, 1} \{\pi(s_i; a_i) + e^{-r\Delta t}E^Q_t[J_{i+1}(I_{i+1}, S_{i+1}, k_{i+1}, \delta')|a_i]|}$$

for $i = 1, 2, \ldots, n$, with $\pi(s_i; a_i)$ given by (2.57), $k_{i+1}$ given by (2.55) and $J_{n+1}(\cdot, \cdot, \cdot, \cdot) \equiv 0$.

The expected present value of $J_1(s_1)$ viewed from date $t_0$ is:

$$V_0(s_0) = e^{-r(t_1-t_0)}E^Q_t[J_1(s_1)]. \quad (2.59)$$

Note that the expectation in (2.58) is taken with respect to the risk neutral pricing measure $Q$. In order to do this, we need the risk neutral process that corresponds to (2.55). This, however, is easy to obtain since the risk that stems from (2.55) is diversifiable. We now continue our analysis with characterization of an optimal policy.

### 2.3.2 A Monotone Optimal Policy

In contrast to the model described in Section 2.2, in general, an optimal policy $a^*_i(s_i)$ is not myopic and a closed form solution similar to that presented in Section 2.2.2 is now lost. However, analysis of the dynamic program (2.58) and the concept of supermodularity allow us to characterize it and gain some general economic insight. This is developed next.
Proposition 15. The value function $J_i(s_i)$ given by (2.58) is non-decreasing in $S_i, k_i,$ and $\delta_i$, and non-increasing in $I_i$ for all $i = 1, 2, \ldots, n$. In addition, $J_i(s_i)$ is non-increasing in $i$.

Proof. See Appendix.

Proposition 15 states that the value of the sequence of options to engage in production at times $t_1 < t_2 < \cdots < t_n$ increases as the price of output increases, as the state of the system improves, and as the cost of input decreases. Moreover, it is also non-decreasing in the number of remaining production epochs.

The next two results utilize the concepts of supermodularity to establish monotonicity properties of an optimal policy; first, with respect to the state vector $s_i$ for a given period, and then with respect to the number of remaining production periods for a given state vector, $s_i$.

Proposition 16. Suppose the cost function $g(k'_i, k_i)$ given by (2.35) is submodular. Then in (2.58), there exists an optimal policy, which is non-decreasing in $S_i, k_i,$ and $\delta_i$, and non-increasing $I_i$, for each $i = 1, 2, \ldots, n$.

Proof. See Appendix.

Proposition 16 asserts that for each time period, there exists an optimal policy under which the profit-maximizing firm is more likely to engage in production and attain higher targeted effort level as the price of output increases, the state of the system, $k$, improves and if the temporary shutdown option has not been exercised. The firm is less likely to engage in production or to attain higher levels of effort as the cost of input material increases.
Since there are only two actions $\delta_i' = \{0, 1\}$ ( = shutdown temporarily or produce), for each time $t_i$, there exists $s_i^* = (I_1^*, S_1^*, k_1^*, \delta_1^*)$ such that if $I_i \geq I_1^*, S_i \leq S_1^*, k_i \leq k_1^*$, and $\delta_i \leq \delta_1^*$, the optimal decision is to shut down temporarily. Similarly, since $k_i'$ is bounded above by 1, there exists $s_i^* = (I_2^*, S_2^*, k_2^*, \delta_2^*)$ such that if $I_i \leq I_2^*, S_i \geq S_2^*, k_i \geq k_2^*$, and $\delta_i \geq \delta_2^*$, the optimal decision is to produce while targeting the maximum production yield. Moreover, $I_1^* \geq I_2^*, S_1^* \leq S_2^*, k_1^* \leq k_2^*$, and $\delta_1^* \leq \delta_2^*$. This observation generalizes the optimal operating policy given in Proposition 10.

The final result examines the relationship between an optimal effort policy and the number of remaining production epochs.

**Proposition 17.** In (2.58), suppose $s_i = s_{i+1} = s$, where $i = 1, 2, \ldots, n - 1$. Then $k_{i+1}^*(s) \leq k_i^*(s)$.

*Proof.* See Appendix.

Everything else being equal, Proposition 17 asserts that as the number of remaining production epochs decreases, a profit-maximizing firm will attain lower optimal targeted effort levels $k_i^*$. Equivalently, a greater number of production epochs warrants higher optimal targeted effort levels $k_i^*$ because more time is available to reap the benefits.

### 2.3.3 An Example

Most problems that satisfy the assumptions made in Section 2.3, require numerical solutions as the temporary shutdown and quality options become complex compound options. For example, see Enders et al. (2007) who study the problem of valuation.
of natural gas wells. To illustrate the results, we consider a problem with \( n = 2 \), \( e_{on} = 0 \), and \( t_0 = t_1 \), conditions under which the solution remains tractable.

Assume that the number of different operating states, \( k \), that can be utilized is \( m \). Without loss of generality, these are labeled \( 0 \leq k^{(1)} < k^{(2)} < \cdots < k^{(m)} \leq 1 \). The feasible values for \( k' \) correspond to the feasible values of \( k \). Committing to a state \( k' = k^{(j')'} \) at the beginning of the period does not ensure that the final state is \( k^{(j')'} \).

Let \( P \) be an \( m \times m \) transition matrix constrained so that the likelihood of being in a high final state is monotone increasing in \( k' \). If no production takes place, then the state of the system remains unchanged.

The expected profit in the last period is:

\[
\pi(I_2, S_2, k_2; k'_2, \delta'_2) = \delta'_2 [k'_2 e^{-qs \Delta t} S_2 - I_2 - g(k'_2, k_2)].
\]  

(2.60)

Let \( k''_2(I_2, S_2, k_2) \) be the \( k^{(j')'} \) for which the right side of (2.60) attains its maximum with \( \delta'_2 = 1 \). That is,

\[
k''_2(I_2, S_2, k_2) = \arg \max_{k^{(j')'}_2} \left\{ k^{(j')'}_2 e^{-qs \Delta t} S_2 - I_2 - g(k^{(j')'}_2, k_2) \right\}.
\]  

(2.61)

With (2.61), the time \( t_2 \) expected profit becomes:

\[
\pi(I_2, S_2, k_2) = \max \left\{ k''_2 e^{-qs \Delta t} S_2 - I_2 - g(k''_2, k_2), 0 \right\}.
\]  

(2.62)

An optimal operating policy at time \( t_2 \) will have a simple structure consistent with the results presented in Section 2.3.2. Continuing with the backward induction, the market value of the option to produce at times \( t_1 \) and \( t_2 \) (viewed from time \( t_1 \)) is
given by:

\[ J_1(I_1, S_1, k_1) = \delta'[k'_1e^{-qs\Delta t}S_1 - I_1 - g(k_1, k'_1)] \]

\[ + \delta'e^{-r\Delta t}E_{t_1}^Q\left[ \sum_{j=1}^{m} p(k'_1, k_2(j))J_2(I_2, S_2, k_2(j)) \right] + (1 - \delta')e^{-r\Delta t}E_{t_1}^Q\left[ J_2(I_2, S_2, k_1) \right], \]  

where \( p(k'_1, k_2(j)) \), an element of the matrix \( P \), denotes a probability of transitioning to the state \( k_2(j) \) given effort \( k'_1 \). With the input price \( I \) constant, the value of \( e^{-r\Delta t}E_{t_1}^Q[J_2(I_2, S_2, k_2(j))] \) obtains in a closed form; otherwise a numerical analysis or an analytical approximation are necessary.

Let \( k'^*_1(I_1, S_1, k_1) \) be the \( k_2(j) \) for which the right side of (2.63) attains its maximum with \( \delta'_1 = 1 \). With \( k'^*_1(I_1, S_1, k_1) \), the right side of (2.63) becomes

\[
\max\left\{ k'^*_1 e^{-qs\Delta t}S_1 - I_1 - g(k_1, k'^*_1) + e^{-r\Delta t}E_{t_1}^Q\left[ \sum_{i=1}^{m} p(k'^*_1, k_2^{(i)})J_2(I_2, S_2, k_2^{(i)}) \right], e^{-r\Delta t}E_{t_1}^Q\left[ J_2(I_2, S_2, k_1) \right] \right\}.
\]

Figure 2.5 illustrates both an optimal operating policy and the monotonicity results from Section 2.3.2. For example, if the system is in state \( k_2 = 0 \) at time \( t_2 \), the price of output \( S_2 \) needs to be at least $34.59 for the firm to be optimal to produce while attaining the highest level of effort. On the other hand, the output price of only $32.07 is needed if the state of the system is \( k_2 = 0 \). This is seen in Panel (a) of Figure 2.5. Finally, Panel (b) of Figure 2.5 illustrates that a market of output of only $22.55 is needed to produce at the highest level of effort if the system is in state \( k_1 = 0 \) at time \( t_1 \).

Using Proposition 13, it is also straightforward to show that the relative contribution of the quality option result holds in this setting. Figure 2.6 illustrates how the
Figure 2.5: Optimal policy at times $t_1$ and $t_2$ as a function of the market price of output $S$ given $q_S = 0.02$, $I = $2, $r = 0.05$, $\sigma_S = 0.45$, $\Delta t = 0.25$ years. There are three states: $k^{(1)} = 0$, $k^{(2)} = 0.5$, $k^{(3)} = 1$. The transition probabilities are: $p(0, 0) = 1$, $p(0.5, 0.5) = 0.625$, $p(0.5, 0) = 1 - p(0.5, 0.5)$, $p(1, 1) = 0.75$, $p(1, 0.5) = 1 - p(1, 1)$, and zero otherwise. The costs are: $g(0, 0) = $0, $g(0, 0.5) = $5.25, $g(0, 1) = $22.375, $g(0.5, 0.5) = $2.75, $g(0.5, 1) = $18.625, $g(1, 1) = $2.5, and infinity otherwise.

Relative value of the quality option declines as price uncertainty increases in a production system with $k_f' = 1$. The numerical result shows that as prices become very uncertain, the value of the quality option is cut nearly in half as the firm increasingly benefits from the simple ability to switch production “on” and “off”.

57
Figure 2.6: Relative value of the quality option at time $t_0$ given $S_0 = $10, $k_0 = 0.5$, $k_f = 1$, $q_S = 0.02$, $I = $2, $r = 0.05$, $\sigma_S = 0.45$, $\Delta t = 0.25$ years, and $t_0 = t_1$. There are three states: $k^{(1)} = 0$, $k^{(2)} = 0.5$, $k^{(3)} = 1$. The transition probabilities are: $p(0,0) = 1$, $p(0.5,0.5) = 0.625$, $p(0.5,0) = 1 - p(0.5,0.5)$, $p(1,1) = 0.75$, $p(1,0.5) = 1 - p(1,1)$, and zero otherwise. The costs are: $g(0,0) = $0, $g(0,0.5) = $5.25, $g(0,1) = $22.375, $g(0.5,0.5) = $2.75, $g(0.5,1) = $18.625, $g(1,1) = $2.5, and infinity otherwise. Note that $\phi_{fix}(S_0,k_0|k_f) = 0.12$, $V_{fix}(S_0,k_0|k_f) = 0.018$, $\phi_0(S_0,k_0) = 2.19$, and $V_0(S_0,k_0) = 0.746$.

2.4 Conclusion

In the context of McDonald and Siegel (1985), yield management in a production system is managed solely by the sequence of temporary shutdown options. We have considered a firm that in addition to the temporary shutdown options has the ability to exert different levels of effort to improve the expected output yield. Modeling this quality option in the presence of possible shutdowns and restarts is what differentiates our study from the classical works of McDonald and Siegel (1985), Brennan and Schwartz (1985), and Dixit and Pindyck (1994).

Our paper provides insights into better understanding conditions whereby the
quality option adds significant value over and above the temporary shutdown option. This analysis may be useful in the design phase of technologies where decisions have to be made about the degree of built in operational flexibility.

Many extensions of our model can be accomplished. First, we have assumed the firm’s inputs and outputs trade in complete and arbitrage free markets. This assumption can easily be weakened. For example, our output prices may only be correlated with prices of financial assets. Second, we have assumed the firm is a price taker. Extending the model to allow for downward sloping demand curves would be of some interest. Burnetas and Ritchken (2005), for example, consider option contracts in supply chains when the demand curve is downward sloping. It remains for future work to study the design issues of operational flexibility in such systems. The greater the array of choices provided by the quality option, the greater its value. However, there is a tradeoff with cost and also with the temporary shutdown option. While our paper provides some fundamental properties of the value of the quality option in the presence of the temporary shutdown option, more work along those lines is called for. Finally, the approach followed here can be readily adapted in the area of stochastic maintenance to derive market sensitive preventive maintenance policies. Similarly, capacity planning analysis and valuation of unreliable production technologies such as power generating assets are well suited for this analysis.
Chapter 3

Risk Averse Supply Chain Coordination

3.1 Introduction

Much of the supply chain contracting literature assumes risk neutrality and analyzes the interaction between the supplier and the retailer as a Stackelberg game; see Cachon (2003), for example, for more details. This popular approach, however, leaves two issues open: (1) it does not specify the payoff allocation, so that it does not provide values for the supplier’s or retailer’s opportunities to enter into a supply contract; and (2) it is not completely obvious how to proceed with the analysis when the supplier and the retailer are risk averse.

This paper presents an alternative game between an upstream supplier and a newsvendor retailer that is meaningfully different from the traditional Stackelberg game. The solution of the game, which utilizes the results presented in seminal papers
by Nash (1950, 1953), sheds some light on the two unaddressed issues. In particular, the Nash solution allocates specific (in many cases unique) monetary payoffs to both players, so that both the supplier and the retailer can “value” the opportunity to enter into a supply contract before the negotiations even begin and gain valuable insight into the strengths of their bargaining positions. Moreover, the Nash solution is amenable to incorporating risk aversion without requiring that the supplier’s or the retailer’s utility functions convey more than *ordinal* information. This feature is particularly attractive because positive affine transformations of either player’s utility function should have no impact on the solution of the game.

The importance of considering risk aversion may not be immediately obvious; after all, for the case of publicly held firms, there are well-known diversification arguments that support the commonly made risk neutrality assumption. Many of them, however, come apart due to factors such as market incompleteness or fail to apply to privately held firms, which may explain why 50 years of research in decision making behavior refutes risk neutrality. Experimental evidence presented in Schweitzer and Cachon (2000), for example, suggests that, for some products, managers are willing to tradeoff lower expected profit for downside protection against possible losses; evidence from practice (cf. Froot et al. (1993)) shows that firms devote substantial resources to risk management. Therefore, it is important to develop supply chain contracting results without assuming risk neutrality.

The idea of incorporating risk sensitivity into supply chain contract research is not new. Some of the existing research extends the Stackelberg approach by using *cardinal* utility functions (e.g. Gan et al. (2004)). Others, such as Choi et al. (2008a,b),
analyze the problem using a mean-variance framework, which is well-known to be equivalent to a quadratic utility function. This approach is analytically tractable and it can be used to obtain approximate solutions; this is due to the results in Levy and Markowitz (1979), who show that some families of utility functions are well approximated by quadratic functions.

Risk aversion tends to decrease the retailer's optimal order quantity (cf. Bouakiz and Sobel (1992), Chen and Fedegruen (2000)). Eeckhoudt et al. (1995), for example, show that a risk averse retailer may order nothing due to high demand uncertainty. To mitigate the otherwise reduced performance of the entire supply chain, Agrawal and Seshadri (2000) explore the role of risk intermediation, which offers another avenue of dealing with aversion to risk, particularly if only one of the players is risk averse.

The applications of bargaining in supply chain, inventory, and capacity management problems include Kohli and Park (1989) who study the implementation of quantity discounts in an inventory problem. Other uses of bargaining models include Reyniers and Tapiero (1995) who apply it to a quality control problem, and Chod and Rudi (2006) who study investments in capacity. Gurnani and Shi (2006) derive a Nash bargaining solution for the problem of an interaction between a retailer and an unreliable supplier when there are asymmetric beliefs in the supplier's ability to deliver.

This paper is most closely related to Nagarajan and Sošić (2008) who motivate the use of the axiomatic Nash bargaining model (cf. Nash (1950)) in supply chain contracting. They evaluate a special case of our model that is a buy back contract when the supplier is risk neutral and the retailer risk averse.
Here, we complement their approach with a more general supply chain bargaining game of complete information in which the supplier and the retailer negotiate order quantity, contract type, and prices. The game meaningfully differs from the Stackelberg game traditionally considered in the supply chain literature. Using the results from Nash (1950, 1953), we present equilibrium supply contracts under risk neutrality and risk aversion, which allow both players to value the game and judge their bargaining strength, namely the two important issues unaddressed by the Stackelberg game.

The next section of this paper introduces the model. Section 3.3 formally describes the game between the supplier and the retailer. Section 3.4 analyzes the game both under risk neutrality and risk aversion. Section 3.5 concludes.

3.2 The Model

The model is the same as in Cachon (2003) with one non-trivial innovation: an upstream supplier sells to a newsvendor retailer who stocks in anticipation of stochastic season demand, \( D \). The supplier and the retailer are expected utility maximizers with utility functions \( U_s(\cdot) \) and \( U_r(\cdot) \). The innovation is that \( U_s(\cdot) \) and \( U_r(\cdot) \) are not necessarily linear, an assumption made in Cachon (2003). We are particularly interested in the cases where \( U_s(\cdot) \), \( U_r(\cdot) \), or both are concave, which is consistent with risk aversion. Both utility functions have only ordinal meaning.

The purpose of the model is to study equilibrium operational policies the supplier and the retailer should follow and the utility they should each expect to achieve by
doing business with each other.

We adopt the following notation: \( F \) is the distribution function of demand (with \( \overline{F} = 1 - F \)); \( p \) is the retail price; \( c_s \) and \( c_r \) are the supplier’s and the retailer’s production costs; \( g_s \) and \( g_r \) are the supplier’s and the retailer’s shortage penalty cost; and \( v \) is the unit net salvage value of unsold goods. Also, let \( c = c_r + c_s \), \( g = g_r + g_s \), and assume \( v < c < p \).

The retailer’s and supplier’s monetary allocations as functions of order quantity \( q \) and demand \( D \) are:

\[
\pi_r = pS + vI - c_r q - g_r L - \sum_{k \in K} i_k T_k \\
= (p - v + g_r)S - (c_r - v)q - g_r D - \sum_{k \in K} i_k T_k, \quad (3.64a)
\]

\[
\pi_s = g_s S - c_s q - g_s D + \sum_{k \in K} i_k T_k, \quad (3.64b)
\]

where \( S = \min\{q, D\} \) denotes sales; \( I = (q - D)^+ \) denotes excess inventory; \( L = (D - q)^+ \) is the lost sales function; \( T_k \) is a transfer payment function; and \( i_k \) is an indicator variable denoting that a particular supply contract is in force. \( K \) is a discrete set of types of contracts.

The transfer payment function \( T_k, k \in K \) represents a monetary transfer from the retailer to the supplier. To merely streamline exposition, we restrict our attention to three well-known supply contracts: buy back, revenue sharing, and quantity discount; this implies \( K = 3 \). The respective transfer payments for these contracts are as
follows:

\[ T_1 = bS + (w_b - b)q, \]  

(3.65a)

Buy back:

\[ T_2 = (w_r + (1 - \phi)v)q + (1 - \phi)(p - v)S, \]  

(3.65b)

Revenue sharing:

\[ T_3 = w_qq, \]  

(3.65c)

Quantity discount:

where \( w \) is the wholesale price, \( b \) is the buy back price, and \( \phi \) is the sales revenue fraction the retailer keeps. Detailed descriptions of each contract are available in Pasternack (1985), Cachon and Lariviere (2005), Tomlin (2003), and Cachon (2003). Since there is a substantial literature available, we skip the details. Equations (3.65), however, are relevant.

3.3 The Game

In order to analyze the retailer’s and the supplier’s optimal policy, the supply chain contracting literature frequently formulates the interaction between the two as a Stackelberg game (cf. Cachon (2003)). The analysis in this paper formulates the same problem as a stylized bargaining game and studies the equilibrium outcomes of the negotiation between the supplier and the retailer.

Formally, let \( M \) be a non-empty, compact set of pure strategies available to both the supplier and to the retailer, and let \( m \) denote an element in \( M \). In particular, \( m \) is a vector that consists of the order quantity, \( q \); a vector of zero-one variables, \( i = (i_1, i_2, i_3) \), where \( i_1 + i_2 + i_3 = 1 \); and a vector of contract parameters, \( t = (w_b, b, w_r, \phi, w_q) \).
Neither player, however, is restricted to pure strategies and may deliberately randomize to choose a pure strategy – this is referred to as a mixed strategy. Let $\Delta(M)$ denote the set of all probability distributions over the set of Borel measurable subsets of $M$. The pure strategy monetary payoffs to the players, given by (3.64), are denoted by $\pi_s(m)$ and $\pi_r(m)$. The shorthand $\pi^\gamma$ and $\pi^\gamma$ is used to denote mixed strategy monetary payoffs; that is

$$\pi^\gamma_j = \int_M \pi_j(m) d\gamma(m), \; \gamma \in \Delta(M), \; j = \{r, s\}. \quad (3.66)$$

These payoffs give rise to expected utility possibility sets

$$U' = \{(EU_s[\pi_s(m)], EU_r[\pi_r(m)]) : m \in M\} \quad \text{and}$$

$$U = \{(EU_s[\pi^\gamma_s), EU_r(\pi^\gamma_r)) : \gamma \in \Delta(M)\}.$$

Both can be easily displayed graphically in the Cartesian plane and for the purpose of our analysis, we assume that they are convex and compact. It is straightforward to show that $U = U'$ when $U_s(\cdot)$ and $U_r(\cdot)$ are both linear functions; otherwise more care is needed, but the equality remains valid. The economic interpretation of the set $U$ is that it is the set of all pairs of expected utility levels that the supplier and the retailer can achieve in this game via strategies $\gamma \in \Delta(M)$.

Similarly to the Stackelberg game, the bargaining problem is a two move game. In the first move, the supplier and the retailer choose some mixed strategies, say $\tau_s$ and $\tau_r$, that they will be forced to use if they cannot agree on a supply contract. The strategies $\tau_s$ and $\tau_r$ are threats used in the contract negotiation; their corresponding monetary payoffs are labeled $\pi^{\tau_s}$ and $\pi^{\tau_r}$. To illustrate, possible threat strategies
may include walking away when faced with zero expected utility or an expected utility level that is less than what is possible to attain by entering into a competing supply contract. Assume both players are informed of each other’s utility functions and threats.

Next, the supplier and the retailer decide independently on the expected utility levels they demand from a potential supply chain contract. We simply label these $EU_s(\delta_s)$ and $EU_r(\delta_r)$, where $\delta_s, \delta_r \in \Delta(M)$. That is, $\delta_s$ and $\delta_r$ are some mixed strategies that reflect the supplier’s and the retailer’s preferred order quantities, prices and contract types.

In the second move of the game, payoffs are determined. In particular, if there exists a mixed strategy $\gamma \in \Delta(M)$ such that $EU_s(\gamma) \geq EU_s(\delta_s)$ and $EU_r(\gamma) \geq EU_r(\delta_r)$, then the supplier and the retailer receive their demanded expected utility levels. Otherwise, threats are carried out and the players attain $(EU_s(\tau_s), EU_r(\tau_r))$. Clearly, given this payoff function, both players have the incentive to demand as much expected utility as possible without losing compatibility, namely $EU_s(\gamma)$ and $EU_r(\gamma)$.

### 3.4 Analysis

Using results from Nash (1950, 1953), every solution $\gamma \in \Delta(M)$ satisfies

$$\max \left[ EU_s(\gamma) - EU_s(\delta_s) \right] \left[ EU_r(\gamma) - EU_r(\delta_r) \right].$$

(3.67)
3.4.1 A Risk Neutral Case

First, we demonstrate the application of the supply contract bargaining approach when \( U_s(\cdot) \) and \( U_r(\cdot) \) are linear (i.e. both players are risk neutral).

Using the Nash bargaining solution (3.67) and the fact that \( U = U' \), the first-order condition with respect to the order quantity \( q \in m \) yields:

\[
\left( E\pi_r^r - E\pi_s^r \right) \frac{\partial E\pi_s^r}{\partial q} + \left( E\pi_s^s - E\pi_s^r \right) \frac{\partial E\pi_r^s}{\partial q} = 0,
\]

which is satisfied whenever

\[
\frac{\partial E\pi_s^r}{\partial q} = \frac{\partial E\pi_r^s}{\partial q} = 0.
\]

Define

\[
\pi(q) = \pi_r(m) + \pi_s(m) = (p - v + g)S - (c - v)q - gD,
\]

where \( \pi_r(m) \) and \( \pi_s(m) \) are given by (3.64), and for some \( 0 \leq \lambda \leq 1 \), let

\[
\begin{align*}
\pi_r(m) &= \lambda \pi(q) + D\lambda g - Dg_r, \\
\pi_s(m) &= (1 - \lambda) \pi(q) - D\lambda g + Dg_r.
\end{align*}
\]

Note that the right side of (3.69) does not depend on the transfer payment function \( T_k, k = 1, 2, 3 \). Using (3.66) and (3.70), (3.68) is equivalent to

\[
\lambda \frac{\partial E\pi(q)}{\partial q} = (1 - \lambda) \frac{\partial E\pi(q)}{\partial q} = 0 \quad \Leftrightarrow \quad \frac{\partial E\pi(q)}{\partial q} = 0
\]

which holds at \( q = q^* \), where

\[
q^* = F^{-1} \left( \frac{p - c + g}{p - v + g} \right).
\]

In addition to the simplification of the first-order condition (3.68), the virtue of the transformation (3.70) is that for a given \( (\lambda, q) \), there exists a unique vector of
parameters \( t = (w_b, b, w_r, \phi, w_q) \) such that the expected values of (3.64a) and (3.70a), and (3.64b) and (3.70b), equal for all \( i = (i_1, i_2, i_3) \), where \( i_1 + i_2 + i_3 = 1 \). Since the details of the mapping between \( \lambda \) and \( t \) are given in Cachon (2003), we don’t include this. However, for illustration, the mapping between the parameter \( \lambda \) and \((w_b, b)\), parameters of the buy back contract, satisfies:

\[
p - v + g_r - b = \lambda(p - v + g) \quad w_b - b + c_r - v = \lambda(c - v)
\]

Equivalent mappings also exist for revenue sharing and quantity discount contracts.

The Nash bargaining solution (3.67) can now be written as

\[
\max_{\lambda \in [0,1]} \left[ (1 - \lambda)E\pi(q) - \mu\lambda g + \mu g_r - E\pi^*_s \right] \left[ \lambda E\pi(q) + \mu\lambda g - \mu g_r - E\pi^*_r \right], \quad (3.73)
\]

where \( \mu \) denotes the expected value of demand, \( D \).

Using (3.73), the first-order condition with respect to \( \lambda \) yields:

\[
(E\pi(q) + \mu g) \left( 2\lambda E\pi(q) - E\pi(q) + E\pi^*_s - E\pi^*_r - 2\mu g_r + 2\lambda \mu g \right) = 0,
\]

which holds at \( \lambda = \lambda^* \), where

\[
\lambda^* = \frac{E\pi(q) + E\pi^*_r - E\pi^*_s + 2\mu g_r}{2(E\pi(q) + \mu g)}. \quad (3.74)
\]

It is now straightforward to show that (3.73) is maximized at the critical point \((q^*, \lambda^*)\) given by (3.72) and (3.74) respectively. Moreover, the equations (3.70), (3.72), and (3.74) now allow us to characterize the Nash solution to the risk neutral supply contract bargaining problem.

**Proposition 18.** In equilibrium, a risk neutral supplier and a risk neutral retailer both demand the order quantity \( q^* \), given by (3.72), that maximizes the expected profit
of the entire supply chain and opt for any contract with a transfer payment function \( T \), given by (3.65), able to allocate \( \lambda^* \) and \( 1 - \lambda^* \) of the total supply chain profit, \( \pi(q^*) \), given by (3.69), to the retailer and to the supplier respectively. The parameter \( \lambda^* \) is given by (3.74).

The standard Stackelberg game analysis as presented in Cachon (2003), for example, yields the same order quantity \( q^* \) as the Nash bargaining solution presented here. The new information is the division of the total supply chain profit, \( \lambda^* \), given by (3.74). Therefore this result confirms the existing economic insight with respect to the optimal order size and sharpens it with respect to the division of the monetary payoffs.

Equation (3.74) also illustrates the role of the threat strategies, \( \tau_r \) and \( \tau_s \), as the right side increases in \( E\pi_r^\tau_r \) and decreases in \( E\pi_s^\tau_s \).

**Corollary 1.** For either player, an expected threat strategy payoff that is higher than that of the opponent translates into capturing a greater portion of the total supply chain profit.

This result highlights how either the supplier or the retailer can leverage competing supply chain contracts in the current negotiation. For the special case when \( g = 0 \) and \( E\pi_r^\tau_r = E\pi_s^\tau_s \), implying that both players are utilizing equally effective threat strategies, the equation (3.74) yields \( \lambda^* = 1/2 \), implying that the total supply chain profit would be split exactly in half.
3.4.2 A Risk Averse Case

In the Stackelberg approach applied to the risk neutral case, the retailer’s optimal order quantity, $q^*$, is set so as to maximize the total expected supply chain profit. In the risk averse setting, however, both the retailer and the supplier are expected utility maximizers, not expected profit maximizers. Maximizing the total expected utility of the supply chain by simply summing the supplier’s and the retailer’s total expected utilities does not work without also requiring that the supplier’s and the retailer’s utility functions have cardinal meaning. Generally, however, utility functions only convey ordinal information. Therefore, it is not immediately obvious how to generalize the analysis when the players are risk averse. This section suggests a different approach by generalizing the results presented in the previous section.

Assume now that the utility functions $U_s(\cdot)$ and $U_r(\cdot)$ are concave and differentiable. As before, the Nash solution to the supply contract bargaining problem is given by (3.67), which, as can be seen, is immune to positive affine transformations of $U_s(\cdot)$ and $U_r(\cdot)$.

The first-order condition with respect to the order quantity, $q$, yields:

$$
\left[EU_r(\pi^*_r) - EU_r(\pi^*_r')\right]\frac{\partial EU_s(\pi^*_s)}{\partial q} + \left[EU_s(\pi^*_s) - EU_s(\pi^*_s')\right]\frac{\partial EU_r(\pi^*_r)}{\partial q} = 0.
$$

Unfortunately, except for a handful of special cases, one can no longer write $\partial EU_s(\pi^*_s)/\partial q$ and $\partial EU_r(\pi^*_r)/\partial q$ as scalar multiples of each other (as in (3.71) by using the transformation (3.70)). Therefore, one no longer has a closed form solution to the bargaining problem as in the risk neutral case.

However, to streamline the exposition and to gain further economic insight into the
problem, we reduce the number of parameters in (3.67) by restricting the permissible supply contract types to buy back and revenue sharing (i.e. for all \( m \in M \) let \( i_3 = 0 \)). The exclusion of the quantity discount contract is a deliberate choice since the quantity discount wholesale price, \( w_q \), that is optimal in (3.67) is many cases too high or too low for a risk averse retailer or a risk averse supplier to agree to it. Therefore the use of the quantity discount contract in this setting is only limited.

Further reduction in the number of parameters is achieved by adopting the following transformation:

\[
(p - v - g_r - b) = \lambda(p - v + g) \quad (w_b - b + c_r - v) = (\lambda - \theta)(c - v) \tag{3.75}
\]

and

\[
\phi(p - v) + g_r = \lambda(p - v + g) \quad w_r + c_r - \phi v = (\lambda - \theta)(c - v), \tag{3.76}
\]

where \( \lambda \in [0, 1] \) and \( \theta \in [\theta_L, \theta_U] \). Note that the size of the interval \([\theta_L, \theta_U]\) is implied by the already existing bounds on the remaining parameters in (3.75) and (3.76); these bounds are a consequence of the compactness of the set \( M \).

Equations (3.75) and (3.76) are one-to-one mappings between the parameters \((w_b, b)\) of the buy back contract and parameters \((w_r, \phi)\) of the revenue sharing contract (cf. equations (3.65)) and parameters \((\lambda, \theta)\).

Let \( \pi_r(m) \), \( \pi_s(m) \), \( \pi(q) \) be given by (3.64a), (3.64b), and (3.69) respectively. Then using (3.75) and (3.76):

\[
\pi_r(m) = \lambda \pi(q) + \theta(c - v)q + D\lambda g - Dg_r, \tag{3.77a}
\]

\[
\pi_s(m) = (1 - \lambda)\pi(q) - \theta(c - v)q - D\lambda g + Dg_r. \tag{3.77b}
\]
As in the risk neutral case, using (3.66), (3.77) allows us to write (3.67) as follows:

$$\max_{\lambda \in [0,1]} \left[ EU_s(\pi_s(m)) - EU_s(\pi_s^r) \right] \left[ EU_r(\pi_r(m)) - EU_r(\pi_r^r) \right],$$  \hspace{1cm} (3.78)

where $\pi_s(m)$ and $\pi_r(m)$ are given by the right side of (3.77).

The structure of the Nash solution to the supply contract bargaining problem with risk averse agents is analogous to that obtained in the risk neutral case.

**Corollary 2.** Let $q^*, \lambda^*$, and $\theta^*$ be optimal in (3.78), then, in equilibrium, a risk averse supplier and a risk averse retailer demand a quantity $q^*$ and opt for either buy back or revenue sharing contract with parameters $(w_b, b, w_r, \phi)$ determined by $(\lambda^*, \theta^*)$ and equations (3.75) and (3.76).

Although, the parameters $q^*, \lambda^*$, and $\theta^*$ cannot generally be given in a closed form, their existence is immediate. This is because, in (3.78), $EU_s(\pi_s(m))$ and $EU_r(\pi_r(m))$ are continuous in $q, \lambda$ and $\theta$ and all three parameters take values on compact intervals. For the case of $q$, note that $q \in m$, $m \in M$, and the set $M$ is compact. The continuity of the expected utilities $EU_s(\pi_s(m))$ and $EU_r(\pi_r(m))$ follows from (3.77) and from differentiability of $U_s(\cdot)$ and $U_r(\cdot)$. Moreover, a closer look at (3.78), yields a contrast between the risk averse solution and the risk neutral one.

First, we compare the risk neutral and the risk averse optimal order quantities $q^*$ and $q^{\ast}$. Using (3.78), the first-order condition with respect to the order quantity, $q$, is given by:

$$\left[ EU_r(\pi_r(m)) - EU_r(\pi_r^r) \right] \frac{\partial EU_s(\pi_s(m))}{\partial q} + \left[ EU_s(\pi_s(m)) - EU_s(\pi_s^r) \right] \frac{\partial EU_r(\pi_r(m))}{\partial q} = 0.$$
Since concavity of $U_s(\cdot)$ and $U_r(\cdot)$ implies $\partial EU_s(\pi_s(m))/\partial q \leq 0$ and $\partial EU_r(\pi_r(m))/\partial q \leq 0$ for all $q \geq q^*$, then we must have $q^* \in [0, q^*]$. This yields the following result.

**Proposition 19.** When compared to the risk neutral order quantity, $q^*$, given by (3.72), the order quantity in a supply chain with risk averse agents is lower.

Next, we focus on the division of the total supply chain monetary payoff, $\pi(q)$. As seen from (3.74), in the risk neutral case, when the supplier and the retailer utilize threat strategies, $\tau_r$ and $\tau_s$, such that $E\pi_{r}^\tau_r = E\pi_{s}^\tau_s$ (and $g = 0$), the total supply chain is split evenly between them. This is no longer the case here: Kannai (1977) concludes that when negotiation takes place between two risk averse agents, then the Nash’s solution assigns a larger share of the total monetary payoff to the less risk averse agent. Roth and Rothblum (1982) and Safra et al. (1990), for example, elaborate by further considering the role of the threat strategies, $\tau_r$ and $\tau_s$. However, in general, being more risk averse than the negotiation counterpart is considered a disadvantage in the Nash’s model.

To conclude, we demonstrate the results presented here on a simple example.

### 3.4.3 An Example

Let $p = 25, c_r = 5, c_s = 10, v = 5$, and $g_r = g_s = 0$. In addition, suppose that the demand is uniformly distributed on $[0, 100]$ and $EU_s(\pi_{s}^\tau_s) = EU_r(\pi_{r}^\tau_r) = 0$.

Using (3.74) and (3.72), the risk neutral order quantity is $q^* = 100($25−$15)/($25−$5) = 50$ and $\lambda^* = 1/2$.

Therefore, in equilibrium, both the supplier and the retailer should demand an
order quantity of 50 units and an even split of the total supply chain profit, \( \pi(q) \). This is achieved, for example, by utilizing the buy back contract with parameters \( w_b = \$15 \) and \( b = \$10 \), meaning that the supplier would charge the retailer a wholesale price of \$15\) and, at the end of the selling season, repurchase all unsold inventory at \$10\) per unit.

Now, suppose \( U_r(x) = x - \alpha x^2 \) and \( U_s(x) = x - \beta x^2 \), where \( \alpha = 0.001 \) and \( \beta = 0.003 \). In this scenario, both the supplier and the retailer are risk averse. Solving (3.78) yields \( q^* = 33, \lambda^* = 0.82 \) and \( \theta^* = -0.14 \). Therefore, as a result of risk aversion, the optimal order quantity decreases from 50 units obtained in the risk neutral case to 33 units. The parameters \( \lambda^*, \theta^* \) determine the equilibrium contract parameters (and hence the division of the monetary payoffs). Using (3.75) and (3.76), they are: \( w_b = \$13.21, b = \$3.60 \) and \( w_r = \$7.30, \phi = \$0.82 \). So that, under the buy back contract, the supplier would charge the retailer \$13.21 per unit and repurchase unsold inventory at \$3.60 per unit. Similarly, under the revenue sharing contract, the supplier would charge the retailer \$7.30 per unit and the retailer would keep 82\% of the sales revenue, passing the remaining 18\% onto the supplier.

More interestingly, suppose now \( \beta = 0 \), so that \( U_s(x) = x \); the retailer’s utility function is still \( U_r(x) = x - 0.001x^2 \). This corresponds to a scenario of a risk neutral supplier supplying a risk averse retailer. This is the case that is considered in Nagarajan and Sošić (2008).

Solving (3.78) in this case yields \( q^* = 50, \lambda^* = 0 \) and \( \theta^* = 0.23 \). This result is interesting because, as it turns out, it is optimal for the risk neutral supplier to
assume the entire supply chain risk. To see this, the parameters $\lambda^* = 0, \theta^* = 0.23$ and (3.77) yield the following monetary payoffs:

$$\pi_r(m) = 0.23(c - v)q$$
$$\pi_s(m) = \pi(q) - 0.23(c - v)q$$

Clearly, the retailer’s monetary payoff is now deterministic, depending only on the order quantity, $q$, and the supplier bears the risk embedded in the component $\pi(q)$ given by (3.69).

The buy back and revenue sharing contracts that achieve these payoffs have parameters $w_b = $17.70, $b = $20 and $w_r = -$7.40, $\phi = 0$.

The revenue sharing contract has an interesting interpretation: $\phi = 0$ indicates that the retailer receives none of the supply chain revenue, however, a compensation of $7.40q$ is provided by the supplier for placing an order of size $q$; this is readily seen from the negative wholesale price, $w_r = -$7.40. Therefore the supplier merely “rents” the shelve space from the retailer.

### 3.5 Conclusion

Utilizing the results in Nash (1950, 1953), this paper presents an alternative method of obtaining equilibrium supply chain contracts, both under risk neutrality and risk aversion. For reasons of exposition, the scope of the paper is limited to three well-known contract types: buy back, revenue sharing, and quantity discount. Moreover, the analysis ignores factors such as the cost of administering these contracts. Relaxing these simplifying assumptions may be of some interest. However, this will increase the analytical complexity; this is true particularly in the risk averse case, where,
in general, closed form solutions do not obtain. We leave these issues for future investigation.
Appendix A

Appendix to Chapter 1

A.1 The Complete Solution

A.1.1 Kharush-Kuhn-Tucker Conditions

In Section 1.5, \((b, s, y) = (b^C(\mu), s^C(\mu), y^C(\mu))\) maximizes \(L(b, s, y, \mu)\) in (1.17) subject to \(b \geq 0\) and \(b + s \geq 0\). The KKT conditions for this problem are as follows:

\[
\frac{\partial L(b, s, y, \mu)}{\partial s} + \lambda_1 = 0 \quad (A.79a)
\]

\[
\frac{\partial L(b, s, y, \mu)}{\partial y} = 0 \quad (A.79b)
\]

\[
\frac{\partial L(b, s, y, \mu)}{\partial b} + \lambda_1 + \lambda_2 = -\rho + \lambda_1 + \lambda_2 = 0 \quad (A.79c)
\]

\[
\lambda_1 (b + s) = 0 \quad (A.79d)
\]

\[
\lambda_2 b = 0 \quad (A.79e)
\]
Equations (A.79a) and (A.79b) are given by

\[
\frac{\partial L(b,s,y,\mu)}{\partial s} = -(1 - \beta) + \begin{cases}
\beta a F\left(\frac{hy-s-\mu(r+h)}{r+h}\right), & \text{if } s + ry > 0; \\
\beta a, & \text{if } s + ry < 0.
\end{cases}
\]

and

\[
\frac{\partial L(b,s,y,\mu)}{\partial y} = -c + \beta r - \beta(r + h - c)F(y - \mu) + \begin{cases}
-\beta ah F\left(\frac{hy-s-\mu(r+h)}{r+h}\right), & \text{if } s + ry > 0; \\
\beta a [r - (r + h)F(y - \mu)], & \text{if } s + ry < 0.
\end{cases}
\]

because the expected value of the penalty function \(p(\cdot)\) given by (1.15) is:

\[
E\{a[s + ry - (r + h)(y - \mu - \varepsilon)^+]\} = \begin{cases}
-a \int_0^{[hy-s-\mu(r+h)]/(r+h)} [hy - s - \mu(r + h)] \, f(x) \, dx, & \text{if } s + ry > 0; \\
-(r + h)x \, f(x) \, dx, & \text{if } s + ry = 0; \\
-a \int_0^{y-\mu} (r + h)(y - \mu - x) \, f(x) \, dx, & \text{if } s + ry < 0; \\
a\{s + ry - (r + h) \int_0^{y-\mu} (y - \mu - x) \, f(x) \, dx\}, & \text{if } s + ry < 0.
\end{cases}
\]

**A.1.2 Complete Solution**

**Lemma 1.** If \((1 - \beta)c \leq \beta r\), an optimal value of \((b, s, y)\) can be determined as follows.

Let \(y_1\) solve:

\[-c + \beta r - \beta(r + h - c)F(y_1 - \mu) - \beta ah F\left(\frac{hy_1 - \mu(r + h)}{r + h}\right) = 0 \quad (A.80)\]

1. If

\[
\frac{1 - \beta}{\beta a} < F\left(\frac{hy_1 - \mu(r + h)}{r + h}\right)
\]

then \(y^C(\mu) = y_0\), \(s^C(\mu) = s_0\) and \(b^C(\mu) = 0\) where

\[
F(y_0 - \mu) = \frac{\beta r - (1 - \beta)h - c}{\beta(r + h - c)}
\]
\[ F \left( \frac{hy_0 - \mu(r + h) - s_0}{r + h} \right) = \frac{1 - \beta}{\beta a} \]

2. If
\[
\frac{1 - \beta - \rho}{\beta a} \leq F \left( \frac{hy_1 - \mu(r + h)}{r + h} \right) \leq \frac{1 - \beta}{\beta a}
\]
then the solution is \( y^C(\mu) = y_1 \) and \( s^C(\mu) = b^C(\mu) = 0 \).

3. If
\[
F \left( \frac{hy_1 - \mu(r + h)}{(r + h)} \right) < \frac{1 - \beta - \rho}{\beta a} \quad (A.81)
\]
then:

(a) If
\[
\frac{1 - \beta - \rho}{\beta a} < \frac{\beta r - (1 - \beta)h - c + h\rho}{\beta(r + h - c)} \quad (A.82)
\]
then \( y^C(\mu) = y_2 \), \( s^C(\mu) = s_2 \) and \( b^C(\mu) = -s_2 \) where
\[
F(y_2 - \mu) = \frac{\beta r - (1 - \beta)h - c + h\rho}{\beta(r + h - c)} \quad (A.83)
\]
\[
F \left( \frac{hy_2 - s_2 - \mu(r + h)}{r + h} \right) = \frac{1 - \beta - \rho}{\beta a} \quad (A.84)
\]

(b) If
\[
\frac{\beta r - (1 - \beta)h - c + h\rho}{\beta(r + h - c)} \leq \frac{1 - \beta - \rho}{\beta a} < 1 \quad (A.85)
\]
then \( y^C(\mu) = y_3 \), \( s^C(\mu) = -ry_3 \) and \( b^C(\mu) = r y_3 \) where
\[
F(y_3 - \mu) = \frac{(1 - \rho)r - c}{\beta a(r + h) + \beta(r + h - c)}
\]

(c) If
\[
1 \leq \frac{1 - \beta - \rho}{\beta a}
\]
then \( y^C(\mu) = y_4 \), \( s^C(\mu) = -\infty \) and \( b^C(\mu) = \infty \) where

\[
F(y_4 - \mu) = \frac{\beta r(1 + a) - c}{\beta a(r + h) + \beta(r + h - c)}
\]

### A.1.3 Additional Comments on Sufficient Conditions

The discussion in the remainder of this paragraph appeals to Lemma 1. Inequalities (1.18) and (1.19) are special cases of (A.80), (A.81) and (A.82) with \( \mu = 0 \). Lemma 2 establishes that (1.19) implies that (A.81) is valid for all \( \mu \geq 0 \). Therefore, \((b^C(\mu), s^C(\mu), y^C(\mu)) = (b_2, s_2, y_2)\) for all \( \mu \geq 0 \).

Moreover, if (A.82) is valid (this is the same as the right-hand inequality in (1.19)) and

\[
\frac{1 - \beta - \rho}{\beta a} < F\left(\frac{hy_1(\mu) - \mu(r + h)}{r + h}\right) \text{ at } \mu = 0
\]

where \( y_1(\mu) \) solves (A.80), then Lemma 3 in Section A.3 implies that (A.81) eventually holds as \( \mu \) increases. The implication of this is that even if the left-hand inequality of (1.19) is invalid, as \( \mu \) increases, all results in the main body of the paper that require (1.19) become valid.

Note that this argument assumes that (A.82) is valid. This condition prevents the optimal solution being \((b^C(\mu), s^C(\mu), y^C(\mu)) = (b_3, s_3, y_3)\) or \((b_4, s_4, y_4)\), which is not of great practical interest. Inspection of Lemma 1 shows that the former requires \( b^C(\mu) = r y^C(\mu) \) implying that default will be triggered by any amount of inventory; and the latter requires \( b^C(\mu) = \infty \) implying that the firm will default with probability one.
A.2 Additional Proofs

Proof. PROPOSITION 5. Confirm (a) and (b) by using (1.17) in $L[(-s)^{+}, s, \mu + f, \mu]$ and applying Leibnitz’s rule to the maximization problems when $s < 0$ and $s \geq 0$. For (c), if $x_n \leq \mu_n + f$ and $y_n = \mu_n + f$ then $x_{n+1} = (y_n - D_n)^{+} = [(\mu_n + f) - (\mu_n + \varepsilon_n)]^{+} = (f - \varepsilon_n)^{+} \leq f \leq \mu_{n+1} + f = y^C(\mu_{n+1})$.

For (d), in the definition of $w_n$ in (1.6b), substitute $z_n = y_n - x_n$, $x_n = (y_{n-1} - D_{n-1})^{+}$, $w_n = g(y_n, D_n) + s_n$, $g(y_{n-1}, D_{n-1}) = ry_{n-1} - (r + h)(y_{n-1} - D_{n-1})^{+}$, $y_n = \mu_n + f$, $y_{n-1} = \mu_{n-1} + f$, and $D_{n-1} = \mu_{n-1} + \varepsilon_{n-1}$, and use $s_{n-1} - s_n = r(\mu_n - \mu_{n-1})$ from (1.22) to obtain

$$v_n = w_n - s_n - p(w_n) - cz_n - \rho b_n$$

$$= g(y_{n-1}, D_{n-1}) + s_{n-1} - s_n - p[g(y_{n-1}, D_{n-1}) + s_{n-1}]$$

$$- cy_n + c(y_{n-1} - D_{n-1})^{+} - \rho b_n$$

$$= rf - (r + h)(f - \varepsilon_{n-1})^{+} + r\mu_n - a(r + h)[(f - \varepsilon_{n-1})^{+} + f_1 - f]^{+}$$

$$- r\rho \mu_n + h\rho f - \rho(r + h)f_1$$

which yields (1.24). □

Proof. PROPOSITION 7. The pairs (1.20), (1.21) and (1.25), (1.27) are the same except that $f_0$ in the former is replaced by $f_3$ in the latter. Since (1.19) ensures $s_n < 0$ in both pairs (for all $n$), a rearrangement of terms yields the following expression that
does not depend on $\mu_1$:

$$
\Delta = E \sum_{n=1}^{\infty} \beta^{n-1} \left[ L\left( [-hf_3 + r\mu + (r+h)f_1]^+, hf_3 - r\mu - (r+h)f_1, \mu_n + f_3, \mu_n \right) 
- L\left( [-hf_0 + r\mu + (r+h)f_1]^+, hf_0 - r\mu - (r+h)f_1, \mu_n + f_0, \mu_n \right) \right] 
= (1 - \beta)^{-1} [\beta(r + h) + \rho h] (f_0 - f_2) - \beta(r + h - c) E \left( [ (f_0 - \epsilon)^+ - (f_2 - \epsilon)^+ \right) 
- \beta a(r + h) E \left( [ (f_0 - \epsilon)^+ - f_0 + f_1]^+ - [(f_2 - \epsilon)^+ - f_2 + f_1]^+ \right)
$$

Proof. PROPOSITION 8. (a) From (1.20) and (1.25), $y^C(\mu) \leq y^D(\mu)$ corresponds to $f_0 \leq f_3$ which is equivalent to $0 \leq \delta$ where

$$
\delta = (r - c)\beta(r + h - c) - (r + h - \beta c)(\beta r - (1 - \beta - \rho)h - c)
$$

Rearranging terms,

$$
\delta = (r - c)\beta(r + h - c) - (r + h - \beta c)(\beta r - c) + h(r + h - \beta c)(1 - \beta - \rho)
\geq (r - c)\beta(r + h - c) - (r + h - \beta c)(\beta r - c) + h(r + h - c)(1 - \beta - \rho)
= (1 - \beta - \rho)h(r + h - c) + c(1 - \beta)[r(1 - \beta) + h] \geq 0
$$

So $y^C(\mu) \leq y^D(\mu)$. The inequality $f_0 \leq f_3$, (1.21), and (1.27) imply $s^C(\mu) \leq s^D(\mu)$.

So $b^C(\mu) \geq b(\mu)$.

(b) Proposition 5. 

83
A.3 Miscellaneous Technical Results

Lemma 2. If (A.82) is valid, \( y_1(\mu) \) solves (A.80) with \( \mu = 0 \), and (A.81) is valid with \( y_1(\mu) \) at \( \mu = 0 \), i.e., if

\[
F\left( \frac{hy_1(0)}{r+h} \right) < \frac{1 - \beta - \rho}{\beta a}
\]

then (A.81) is valid for all \( \mu > 0 \).

Proof. Let \( y_1(\mu) \) solve (A.80) for some \( \mu > 0 \). Since \( F(\cdot) \) is nondecreasing, it is sufficient to show

\[
\frac{hy_1(0)}{r+h} \geq \frac{hy_1(\mu) - \mu(r+h)}{r+h}
\]

By contradiction, suppose

\[
\frac{hy_1(0)}{r+h} < \frac{hy_1(\mu) - \mu(r+h)}{r+h} \tag{A.86}
\]

So

\[
y_1(0) < y_1(\mu) - \frac{\mu(r+h)}{h}
\]

Assuming \( 1 < (r+h)/h \) and \( y_1(0) \geq 0 \),

\[
y_1(\mu) - \mu > y_1(\mu) - \frac{\mu(r+h)}{h} \quad \text{and} \quad y_1(\mu) - \mu > y_1(0) \tag{A.87}
\]

However, (A.86) and (A.87) imply

\[
0 = -(1 - \beta)c + \beta r - \beta(r+h)F(y_1(0)) - \beta ahF\left( \frac{hy_1(0)}{r+h} \right)
\]

\[
\geq -(1 - \beta)c + \beta r - \beta(r+h)F(y_1(\mu) - \mu) - \beta ahF\left( \frac{hy_1(\mu) - \mu(r+h)}{r+h} \right) \tag{A.88}
\]

(with strict inequality if \( F(\cdot) \) is strictly increasing at either argument of \( F(\cdot) \) on the right side of (A.88)) contradicting the fact that \( y_1(\mu) \) solves (A.80) for some \( \mu > 0 \). \( \square \)
Lemma 3. If (A.82) is valid, $F(\cdot)$ is strictly increasing at $[hy_1(\mu) - \mu(r + h)]/(r + h)$, $y = y_1(\mu)$ is a solution of (A.80) at $\mu = 0$, and

$$F\left(\frac{hy_1(\mu) - \mu(r + h)}{r + h}\right) > \frac{1 - \beta - \rho}{\beta a} \text{ at } \mu = 0$$

then there exists $\mu^* > 0$ such that

$$F\left(\frac{hy_1(\mu) - \mu(r + h)}{r + h}\right) < \frac{1 - \beta - \rho}{\beta a}, \forall \mu \geq \mu^*$$

Proof. Let $\mu_1 = 0$ in Lemma 4 and use the fact that $F(\cdot)$ is strictly increasing. \qed

Lemma 4. If (A.82) is valid, $F(\cdot)$ is strictly increasing either at $[hy_1(\mu_2) - \mu_2(r + h)]/(r + h)$ or at $(y_1(\mu_2) - \mu_2)$, and $y_1(\mu_1)$ and $y_1(\mu_2)$ are solutions of (A.80) where $0 \leq \mu_1 < \mu_2$, then

$$\frac{hy_1(\mu_1) - \mu_1(r + h)}{r + h} > \frac{hy_1(\mu_2) - \mu_2(r + h)}{r + h}$$

Proof. By contradiction, if

$$\frac{hy_1(\mu_1) - \mu_1(r + h)}{r + h} \leq \frac{hy_1(\mu_2) - \mu_2(r + h)}{r + h} \quad (A.89)$$

then (A.89) implies

$$h[y_1(\mu_1) - \mu_1] \leq h[y_1(\mu_2) - \mu_2] - r(\mu_2 - \mu_1) < h[y_1(\mu_2) - \mu_2]$$

or

$$y_1(\mu_2) - \mu_2 < y_1(\mu_2) - \mu_2 \quad (A.90)$$

But (A.89) and (A.90) imply that both $y_1(\mu_1)$ and $y_1(\mu_2)$ cannot solve (A.80). \qed
Appendix B

Appendix to Chapter 2

B.1 Proofs

Proof. Proposition 10. Let $S_1^* = 2e^{qS\Delta t}\sqrt{cx_i}$, $S_2^* = 2ce^{qS\Delta t}$, $S_3^* = e^{qS\Delta t}(c + x_i)$, and $x_i = K + I_i$.

Since $g(k')$ is given by (2.40), then the time $t_i$ expected profit $\pi(I_i, S_i)$ is given by (2.42). From (2.42),

$$\max \left\{ \frac{e^{-2qs\Delta t} S_i^2}{4c} - X_i, 0 \right\} = \max \{ e^{-qs\Delta t} S_i - X_i - c, 0 \} = 0$$

(B.91)

for all $S_i < S_1^*$. Hence the firm would never commit to production at time $t_i$ if $S_i < S_1^*$.

The equality in (B.91) follows because

$$\frac{e^{-2qs\Delta t} S_i^2}{4c} - X_i \geq e^{-qs\Delta t} S_i - X_i - c$$

(B.92)

and for all $S_i < S_1^*$

$$\frac{e^{-2qs\Delta t} S_i^2}{4c} - X_i < 0.$$
The inequality (B.92) follows because

$$
\left(\frac{e^{-2qs\Delta t}S_i^2}{4c} - X_i\right) - \left(e^{-qs\Delta t}S_i - X_i - c\right) = \frac{e^{-2qs\Delta t}(S_i - 2cs^q\Delta t)^2}{4c} \geq 0.
$$

(i) Assume $X_i < c$. Then $S_1^* < S_2^*$. This is because

$$
X_i < c \iff 2e^{qs\Delta t} \sqrt{cX_i} < 2e^{qs\Delta t}c \iff S_1^* < S_2^*.
$$

By (B.91), $\delta^* = 0$ for all $S_i < S_1^*$ and in particular $\delta^* = 0$ if $X_i < c$. This proves (2.43a).

If $S_1^* \leq S_i \leq S_2^*$, then, by (B.91), $\delta^* = 1$ and by the right side of (2.41), the firm’s optimal effort $k^*$ is $0 < k^* = (e^{-qs\Delta t}S_i)/2c < 1$. This proves (2.43b).

If $S_2^* \leq S_i$, then, by (B.93), also $S_1^* < S_i$. By (B.91), the firm would commit to production at time $t_i$, so that $\delta^* = 1$. By the right side of (2.41), $k^* = 1$, so that the firm would attain the highest level of effort. This proves (2.43c).

(ii) Assume $X_i \geq c$. Then $S_1^* \geq S_2^*$. This is because

$$
X_i \geq c \iff 2e^{qs\Delta t} \sqrt{cX_i} \geq 2e^{qs\Delta t}c \iff S_1^* \geq S_2^*.
$$

The firm’s expected profit at time $t_i$ is in this case only given by

$$
\max\{e^{-qs\Delta t}S_i - X_i - c, 0\}.
$$

This follows immediately from (2.42) because, by (B.91), the firm would never commit to production if $S_i < S_1^*$ and, by (B.94), $S_1^* \geq S_2^*$. However, for all $S_i < S_3^*$, (B.95) implies

$$
\max\{e^{-qs\Delta t}S_i - I_i - K - c, 0\} = 0.
$$

87
Therefore the firm would never commit to production at time \( t_i \) if \( X_i \geq c \) and \( S_i < S_3^* \). This proves (2.44a).

If \( S_i \geq S_3^* \), then (B.96) immediately yields \( \delta^{t*} = 1 \). Since \( X_i \geq c \), then also \( S_3^* \geq S_2^* \) and from the right side of (2.41), \( k^{t*} = 1 \). The inequality \( S_3^* \geq S_2^* \) follows because

\[
X_i \geq c \Leftrightarrow e^{qS\Delta t}(c + X_i) \geq 2e^{qS\Delta t}c \Leftrightarrow S_3^* \geq S_2^*.
\]

This completes the proof of (2.44b).

Proof. PROPOSITION 11. Using (2.42), the expression for \( v(i)(I_0, S_0) \) in (2.45b) can be written as:

\[
v(i)(I_0, S_0) = e^{-r(t_i-t_0)} \left\{ E^Q_{t_0} \left[ \max \left\{ \frac{e^{-2qs\Delta t}S^2_i}{4c} - I_i - K, 0 \right\} \left| S_i < e^{qs\Delta t}2c \right. \right] \\
+ E^Q_{t_0} \left[ \max \left\{ e^{-qs\Delta t}S_i - I_i - K - c, 0 \right\} \left| S_i \geq e^{qs\Delta t}2c \right. \right] \\
P\left\{ S_i < e^{qs\Delta t}2c \right\} + E^Q_{t_0} \left[ \max \left\{ e^{-qs\Delta t}S_i - I_i - K - c, 0 \right\} \left| S_i \geq e^{qs\Delta t}2c \right. \right] \\
P\left\{ S_i \geq e^{qs\Delta t}2c \right\} \right\}.
\]

Let \( n(\cdot, \mu, \sigma) \) denote the probability density function of the normal distribution with mean \( \mu \) and standard deviation \( \sigma \) and define \( \mu(i) = \log(S_0) + (r - q_S + \sigma_S/2)(t_i - t_0) \).

(i) If \( X < c \), then using Proposition 10 part (i) and the fact that \( S_i \) follows (2.33b), (B.97) can be written as

\[
v(i)(S_0) = e^{-r(t_i-t_0)} \int_{\log(2e^{qS\Delta t}c)}^{\log(2e^{qS\Delta t}e^{2x}cX)} \left( \frac{e^{-2qs\Delta t}e^{2x}}{4c} - X \right) n(x; \mu(i), \sigma_S\sqrt{t_i-t_0}) dx \\
+ e^{-r(t_i-t_0)} \int_{\log(2e^{qS\Delta t}e^{2x}cX)}^{\infty} \left( e^{-qs\Delta t}e^{x} - X - c \right) n(x; \mu(i), \sigma_S\sqrt{t_i-t_0}) dx.
\]
(ii) If $X \geq c$, then using Proposition 10 part (ii) and the fact that $S_i$ follows (2.33b), (B.97) can be written as

$$v^{(i)}(S_0) = e^{-r(t_i-t_0)} \int_{\log[e^{qS_0}]}^{\infty} \left( e^{-qS_0 e^x - X - c} n(x; \mu^{(i)}, \sigma_S \sqrt{t_i-t_0}) dx. \right.$$  

(B.99)

Simplification of (B.98) and (B.99) yields (2.46) and (2.47) respectively.

\[\square\]

**Proof. Proposition 13.** The inequality (2.50), $v^{(i)}_f(S_0; \sigma_S | k'_f) > 0$, $v^{(i)}_0(S_0; \sigma_S) > 0$, $\phi^{(i)}_f(S_0 | k'_f) > 0$ and $\phi^{(i)}_0(S_0) > 0$ are the sufficient conditions for (2.53a) to hold.

By Corollary 3 in Section B.2, (2.50) does not hold at $k'_f = \sqrt{(K+I)/c}$ and holds with strict inequality for all $k'_f \to 0, 1$. This and the fact that $\phi^{(i)}_f(S_0 | k'_f)/v^{(i)}_f(S_0 | k'_f)$ is continuous and strictly increasing in $k'_f$ on the interval $(\sqrt{(K+I)/c}, 1]$ (cf. Lemma 5 in Section B.2) immediately yields that there exists a unique $k'^{(i)}_U \in (\sqrt{(K+I)/c}, 1)$ such that

$$\frac{\phi^{(i)}_f(S_0; \sigma_S | k'_f = k'^{(i)}_U)}{v^{(i)}_f(S_0; \sigma_S | k'_f = k'^{(i)}_U)} = \frac{\phi^{(i)}_0(S_0; \sigma_S)}{v^{(i)}_0(S_0; \sigma_S)}.$$

Then (2.50) holds for all $k'_f \in (k'^{(i)}_U, 1]$.

By the same argument, there exists $k'^{(i)}_L \in (0, \sqrt{(K+I)/c})$ such that

$$\frac{\phi^{(i)}_f(S_0; \sigma_S | k'_f = k'^{(i)}_L)}{v^{(i)}_f(S_0; \sigma_S | k'_f = k'^{(i)}_L)} = \frac{\phi^{(i)}_0(S_0; \sigma_S)}{v^{(i)}_0(S_0; \sigma_S)}.$$

Moreover, $k'^{(i)}_L$ is unique because $\phi^{(i)}_f(S_0; \sigma_S | k'_f)/v^{(i)}_f(S_0; \sigma_S | k'_f)$ is strictly decreasing in $k'_f \in (0, \sqrt{(K+I)/c})$ (cf. Lemma 5). Then (2.50) also holds for all $k'_f \in (0, k'^{(i)}_L)$. This proves (2.53a). The proof of (2.53b) is analogous. \[\square\]

**Proof. Proposition 14.** Use Proposition 13.

---

89
Definition 1. Let \( I_i^0 = -I_i, s'_i = \left( I_i^0, S_i, k_i, \delta_i \right), s'^+_i = \left( I_i^{p_+}, S_i^{p_+}, k_i^{p_+}, \delta_i^{p_+} \right), s'^-_i = \left( I_i^{p_-}, S_i^{p_-}, k_i^{p_-}, \delta_i^{p_-} \right), \) where \( I_i \geq 0, S_i \geq 0, k_i, k_i' \in [0, 1], \) and \( \delta_i, \delta_i' = \{0, 1\}. \)

We write \( s'^+_i \geq s'^-_i \) if \( I_i^{p_+} \geq I_i^{p_-}, S_i^{p_+} \geq S_i^{p_-}, k_i^{p_+} \geq k_i^{p_-} \) and \( \delta_i^{p_+} \geq \delta_i^{p_-} \). Similarly, \( a'^+_i \geq a'^-_i \) if \( \delta_i^{p_+} \geq \delta_i^{p_-} \) and \( k_i^{p_+} \geq k_i^{p_-} \).

Proof. PROPOSITION 15. Suppose the supremum in (2.58) is attained. Then \( \pi(s'_i; a'^*_i) \) is non-decreasing in \( s'_i \). Due to (2.56) and (2.55), the probability that the state \( k_{i+1} \) in the decision epoch \( t_{i+1} \) exceeds some constant in \([0, 1]\) is non-decreasing in \( k_i \) for all \( k_i'^* \) and \( \delta_i'^* \).

The first part of the proposition is proved by backward induction on \( i \) where \( i = 1, 2, \ldots, n+1 \) starting with \( J_{n+1}(s'_{n+1}) \equiv 0 \), which is also non-decreasing in \( s'_{n+1} \). The assertion regarding monotonicity of the value function \( J_i(s_i) \) with respect to \( i \) is a consequence of the fact that it is a sum of the functions \( \pi_i(s'_i; a'^*_i) \), which are bounded below by zero.

Proof. PROPOSITION 16. \( s'_i \) is a non-empty partially ordered set, where the ordering is specified in Definition 1 and \( (s'_i, a_i) \) is a lattice in the sense of Definition 8-1 in Heyman and Sobel (2004). The result is now proved due to Theorem 8-4 in Heyman and Sobel (2004) and Lemma 6 in Section B.2.

Proof. PROPOSITION 17. Result follows from supermodularity of \( J_j(s'_j) \) on \( j \times k'_j \), where \( j = n - i + 1 \) (i.e. \( j \) represents the number of production periods remaining) and from Theorem 8-4 in Heyman and Sobel (2004). The proof of supermodularity, similar to that of Lemma 6 in Section B.2.
B.2 Lemmas

Lemma 5. Assumptions regarding $c, I, K,$ and $S$ are the same as in Proposition 13. Then $\phi_{fix}^{(i)}(S_0|k_f')/v_{fix}^{(i)}(S_0|k_f')$ is strictly decreasing for all $k_f' \in (0, \sqrt{(K+I)/c})$ and strictly increasing for all $k_f' \in (\sqrt{(K+I)/c}, 1]$, where $v_{fix}^{(i)}(S_0|k_f')$ and $\phi_{fix}^{(i)}(S_0|k_f')$ are given by (2.48) and (2.52) respectively.

Proof. Lemma 5. Define $X = K+I$, $\phi_{fix}' = \partial\phi_{fix}^{(i)}(S_0|k_f')/\partial k_f'$, $v_{fix}' = \partial v_{fix}^{(i)}(S_0|k_f')/\partial k_f'$, and $\Upsilon^{(i)}(k_f') = \phi_{fix}^{(i)}(S_0|k_f')/v_{fix}^{(i)}(S_0|k_f')$. Then

$$\frac{\partial \Upsilon^{(i)}(k_f')}{\partial k_f'} = \frac{\phi_{fix}'(S_0|k_f') - \phi_{fix}^{(i)}(S_0|k_f') v_{fix}'}{v_{fix}^{(i)}(S_0|k_f')^2}, \quad (B.100)$$

where

$$\phi_{fix}' = S_0 e^{-qs(t_i-t_0+\Delta t)n(d^{(i)}_1)} \left( 1 - k_f' d^{(i)}_1 \frac{\partial d^{(i)}_1}{\partial k_f'} \right) \sqrt{t_i-t_0}, \quad (B.101a)$$
$$v_{fix}' = e^{-qs(t_i-t_0+\Delta t)} S_0 N(d^{(i)}_1) - e^{-r(t_i-t_0)} 2ck_f' N(d^{(i)}_2). \quad (B.101b)$$

In (B.100), substitute for $\phi_{fix}'$ from (B.101a) and for $\phi_{fix}^{(i)}(S_0|k_f')$ from (2.52) to obtain

$$\frac{\partial \Upsilon^{(i)}(k_f')}{\partial k_f'} = \frac{S_0 e^{-qs(t_i-t_0+\Delta t)n(d^{(i)}_1)} \sqrt{t_i-t_0} \left[ 1 - k_f' d^{(i)}_1 \frac{\partial d^{(i)}_1}{\partial k_f'} \right] v_{fix}^{(i)}(S_0, k_f') - k_f' v_{fix}' }{v_{fix}^{(i)}(S_0|k_f')^2}$$

from which it immediately follows that

$$\frac{\partial \Upsilon^{(i)}(k_f')}{\partial k_f'} \left( \leq 0 \Rightarrow \left[ 1 - k_f' d^{(i)}_1 \frac{\partial d^{(i)}_1}{\partial k_f'} \right] v_{fix}^{(i)}(S_0, k_f') - k_f' v_{fix}' \right) \left( \leq 0 \right) \quad (B.102)$$

Next, in the right side of (B.102), substitute for $v_{fix}^{(i)}(S_0, k_f')$ and $v_{fix}'$ from (2.48) and (B.101b) respectively and use the fact that

$$k_f' \frac{\partial d^{(i)}_1}{\partial k_f'} = \frac{X - ck_f^2}{\sigma_S(X + ck_f^2) \sqrt{t_i-t_0}}$$
to obtain

$$\frac{\partial \Upsilon(i)(k'_f)}{\partial k'_f} \underset{\langle \rangle}{\underset{\langle \rangle}{\rangle}} 0 \Leftrightarrow \frac{(ck'_f - X)C}{\sigma_S \sqrt{t_i - t_0}} \underset{\langle \rangle}{\underset{\langle \rangle}{\rangle}} 0 \Leftrightarrow (ck'_f - X) \underset{\langle \rangle}{\underset{\langle \rangle}{\rangle}} 0 \Leftrightarrow k'_f \underset{\langle \rangle}{\underset{\langle \rangle}{\rangle}} \sqrt{\frac{X}{c}}, \quad (B.103)$$

where

$$C = k'_f S_0 e^{-q_S(t_i-t_0+\Delta t)} N(d_1^{(i)})d_1^{(i)} - e^{-r(t_i-t_0)}(ck'_f + X)(1 - \sigma_S \sqrt{t_i - t_0}) N(d_2^{(i)})d_2^{(i)} \frac{ck'_f + X}{ck'_f + X}.$$  

A simple argument shows that $C$ is positive for all $k'_f \in (0, \infty)$. Finally, from (B.103):

$$\frac{\partial \Upsilon(i)(k'_f)}{\partial k'_f} \underset{\langle \rangle}{\underset{\langle \rangle}{\rangle}} 0 \Leftrightarrow k'_f \underset{\langle \rangle}{\underset{\langle \rangle}{\rangle}} \sqrt{\frac{X}{c}} \text{ and } k'_f \in (0, 1].$$

Corollary 3. Assumptions regarding $c, I, K,$ and $S$ are the same as in Proposition 13. Then

(i)

$$\phi^{(i)}_{\text{fix}}(S_0; \sigma_S | k'_f) \underset{\langle \rangle}{\underset{\langle \rangle}{\rangle}} \phi^{(i)}_0(S_0; \sigma_S), \quad k'_f = \sqrt{\frac{K + I}{c}};$$

(ii)

$$\frac{\phi^{(i)}_{\text{fix}}(S_0; \sigma_S | k'_f)}{\phi^{(i)}_{\text{fix}}(S_0; \sigma_S | k'_f)} > \frac{\phi^{(i)}_0(S_0; \sigma_S)}{\phi^{(i)}_0(S_0; \sigma_S)}, \quad k'_f \to 0, 1;$$

where $\phi^{(i)}_0(S_0)$ and $\phi^{(i)}_{\text{fix}}(S_0 | k'_f)$ are given by (2.46) and (2.48) respectively; $\phi^{(i)}_0(S_0)$ and $\phi^{(i)}_{\text{fix}}(S_0 | k'_f)$ are the first derivatives of $\phi^{(i)}_0(S_0)$ and $\phi^{(i)}_{\text{fix}}(S_0 | k'_f)$ with respect to $\sigma_S$ and are given by (2.51) and (2.52) respectively.

Proof. Only part (i) is included; the proof of part (ii) is similar. First, note that $\phi^{(i)}_{\text{fix}}(S_0 | k'_f)$ and $\phi^{(i)}_0(S_0)$ are closely related: in the former, the effort level is fixed
at $k'_f$; in the latter the effort level $k'$ is a random variable given by the right side of (2.41). By Lemma 5 in Section B.2, the value of $\phi^{(i)}_{fix}(S_0; \sigma_S | k'_f)/v^{(i)}_{fix}(S_0; \sigma_S | k'_f)$ is uniquely minimized on the interval $(0, 1]$ at $k'_f = \sqrt{(K + I)/c}$. Since the effort level $k'$ in $\phi^{(i)}_f(S_0; \sigma_S)/v^{(i)}_f(S_0; \sigma_S)$ is not fixed at $\sqrt{(K + I)/c}$, the inequality in (i) is immediate.

**Lemma 6.** In the dynamic program (2.58), the function

$$\pi(s'_i; a_i) + e^{-r\Delta t}E_{t_i}^{Q}[J_{i+1}(s'_{i+1})|(s'_i; a_i)]$$

is weakly supermodular.

**Proof.** Using Definition 3 and Lemma 7 in Section B.2, the functions $\delta'_i k'_i e^{-q_S \Delta t} S_i; \delta'_i I_i$ and $-c_{on}(\delta'_i - \delta_i)^+$ are all weakly supermodular on $s'_i \times a_i$. The function $-\delta'_i g(k_i, k'_i)$ is supermodular on $k_i \times k'_i$ because the function $g(k_i, k'_i)$ is assumed to be submodular. Supermodularity on $\delta'_i \times k_i$ is proved using Definition 3: $-g(k^+_i, k'_i) \geq -g(k^-_i, k'_i) \Rightarrow g(k^+_i, k'_i) \leq g(k^-_i, k'_i)$, which holds by assumptions made in Section 2.3. Therefore $\pi(s'_i; a_i)$ is weakly supermodular. This follows because sums of weakly supermodular functions are weakly supermodular.

Weak supermodularity of the function $E_{t_i}^{Q}[J_{i+1}(s'_{i+1})|(s'_i; a_i)]$ on $s'_i \times a_i$ is immediate once supermodularity on $k_i \times \delta'_i$ is established. We must show:

$$E_{t_i}^{Q}[J_{i+1}(s'_{i+1})|(s^{\oplus}_i; k', 1)] - E_{t_i}^{Q}[J_{i+1}(s'_{i+1})|(s^{\ominus}_i; k', 0)]$$

$$\geq E_{t_i}^{Q}[J_{i+1}(s'_{i+1})|(s^{\oplus}_i; k', 1)] - E_{t_i}^{Q}[J_{i+1}(s'_{i+1})|(s^{\ominus}_i; k', 0)],$$

where $s^{\oplus}_i \equiv (-I_i, S_i, k^+_i, \delta_i), s^{\ominus}_i \equiv (-I_i, S_i, k^-_i, \delta_i)$ for all $k^+_i \geq k^-_i$. However, (B.104) holds due to assumption (2.56) and the fact that $J_{i+1}(s'_{i+1})$ is non-decreasing in $s'_{i+1}$.
for all \( i = 1, 2, \ldots, n \) (cf. Proposition 15).

**Definition 2.** Let \( x^+ = (x_1^+, x_2^+, \ldots, x_m^+) \), \( x^- = (x_1^-, x_2^-, \ldots, x_m^-) \), \( y^+ = (y_1^+, y_2^+, \ldots, y_n^+) \), and \( y^- = (y_1^-, y_2^-, \ldots, y_n^-) \), where \( n, m \in \mathbb{N} \). We write \( x^+ \geq x^- \) and \( y^+ \geq y^- \) if \( x_i^+ \geq x_i^- \) and \( y_i^+ \geq y_i^- \) for all \( i = 1, 2, \ldots, m \) and \( j = 1, 2, \ldots, n \).

**Definition 3.** Suppose \( x^+ \geq x^- \), \( y^+ \geq y^- \) (cf. Definition 2). We say that a function \( g(\cdot, \cdot) \) is weakly supermodular if

\[
 g(x^+, y^+) - g(x^+, y^-) \geq g(x^-, y^+) - g(x^-, y^-). \tag{B.105}
\]

If \( m = n = 1 \) and (B.105) holds, then we say a function is supermodular in \( x \times y \).

**Lemma 7.** If \( \partial^2 g(x, y)/\partial x_i \partial y_j \) exists for all \( i = 1, 2, \ldots, m \) and \( j = 1, 2, \ldots, n \), then \( g(x, y) \) is weakly supermodular (cf. Definition 3) if and only if

\[
 \frac{\partial^2 g(x, y)}{\partial x_i \partial y_j} \geq 0
\]

for all \( i = 1, 2, \ldots, m \) and \( j = 1, 2, \ldots, n \).

**Proof.** "\( \Leftarrow \)" If \( g(x, y) \) is weakly supermodular, then by (B.105):

\[
 g(x^+, y^+) - g(x^+, y^-) \geq g(x^-, y^+) - g(x^-, y^-)
\]

for all \( x^+ \geq x^- \) and \( y^+ \geq y^- \). Now let \( y_j^+ \) converge to \( y_j^- \) for all \( j = 1, 2, \ldots, n \) to obtain

\[
 \sum_{j=1}^n \frac{\partial g(x^+, y)}{\partial y_j} dy_j \geq \sum_{j=1}^n \frac{g(x^-, y)}{\partial y_j} dy_j \quad \text{when } x^+ \geq x^- 
\]

implying that

\[
 \sum_{i=1}^m \sum_{j=1}^n \frac{\partial^2 g(x, y)}{\partial y_j \partial x_i} dx_i dy_j \geq 0.
\]
⇒ If \( \frac{\partial^2 g(x, y)}{\partial x_i \partial y_j} \geq 0 \) for \( i = 1, 2, \ldots, m \) and \( j = 1, 2, \ldots, n \), then for \( x^+ \geq x^- \) and \( y^+ \geq y^- \),

\[
\int_{x^-}^{x^+} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\partial^2 g(x, y)}{\partial y_j \partial x_i} \, dx_i \, dy_j \geq 0
\]

\[
\int_{y^-}^{y^+} \sum_{j=1}^{n} \frac{\partial g(x^+, y)}{\partial y_j} \, dy_j - \int_{y^-}^{y^+} \sum_{j=1}^{n} \frac{\partial g(x^-, y)}{\partial y_j} \, dy_j \geq 0
\]

or

\[
g(x^+, y^+) - g(x^-, y^-) \geq g(x^-, y^+) - g(x^-, y^-),
\]

which shows that \( g(x, y) \) is weakly supermodular.
Bibliography


