Technical Memorandum Number 825

Optimization of Inventory and Dividends with Risky Debt

by

Qiaohai (Joice) Hu
Krannert School of Management, Purdue University, hu23@purdue.edu

Matthew J. Sobel
Weatherhead School of Management, Case Western Reserve University,
mjs13@case.edu,

Danko Turcic
Olin School of Business, Washington University in St. Louis, turcic@wustl.edu

May 2010
Optimization of Inventory and Dividends with Risky Debt

March 20, 2010

Qiaohai (Joice) Hu
Krannert School of Management, Purdue University, hu23@purdue.edu

Matthew J. Sobel
Weatherhead School of Management, Case Western Reserve University, mjs13@case.edu

Danko Turcic
Olin School of Business, Washington University in St. Louis, turcic@wustl.edu

Most inventory in the U.S. is purchased on credit and approximately one-third of credit purchases use bank credit. Bank credit purchases tend to be more complex than their trade credit counterparts because firms have more leeway in the choice of maturity and debt size. The aftermath of a default on bank credit is also more complicated than the aftermath of a default on trade credit. Under the assumption that the equity holders would relinquish their claim to future earnings of the firm if bankruptcy were to occur, we study a dynamic stochastic model of inventory and financial decisions, and compare the optimal coordination of the short-run decisions with the corresponding decisions in a decentralized firm. Regardless of whether the capital structure is all-equity or includes long-term debt, the inventory levels, short-term loans, and market value of the equity are higher when inventory and financial flows are coordinated. We also show that the coordinated firm should be more highly leveraged (in terms of long-term debt) than if it were decentralized. At first, the results are obtained under the assumption that product demands are independent and identically distributed random variables. Then we find sufficient conditions for the results to remain valid with non-stationary demand processes.

Key words: Dynamic, inventory, bankruptcy, financing, non-stationary demand, long-term debt

History:

1. Introduction
1.1. Background

U.S. firms rely heavily on credit to finance inventory. Trade credit is the single most important source of inventory financing (cf. Peterson and Rajan (1997)). However, firms are also commonly financed by financial institutions rather than by suppliers. Approximately two-thirds of total liabilities in a typical U.S. firm are represented by trade credit and the remaining one-third by bank credit (cf. Rajan and Zingales (1994)). Lately, Enslow (July, 2008) observes that bank credit has been gaining in importance as some suppliers “...have much less access to credit ... in some cases
may no longer have enough cash to fund orders fully ...” Demica (2007) research report cites lines of credit from a relationship bank as the area with the highest expected growth rate among common finance techniques.

There are important differences between the two types of credit. With trade credit, loan size and loan maturity are essentially predetermined: the loan size is wholesale price times inventory position; and the firm enters the trade credit loan contract when the order is received and discharges it by remitting a portion of the sales revenue to the supplier/lender (cf. Gupta and Wang (2009)). Therefore, with trade credit, financial and material flows are automatically coordinated.

In contrast, when production (or procurement) is financed with bank credit, the loan size does not necessarily equal the total cost of inventory because firms often simultaneously borrow, pay dividends, and cover other operational costs (cf. Brander and Lewis (1988), p. 225). The timing of inventory and borrowing decisions is usually not synchronized either: in large firms, cash and material flows are often managed independently in separate “silos.” In practice, firms typically first secure a credit line, then make an inventory decision that does not necessarily consider financing, and finally use the available credit line to cover the difference between the available cash and the total inventory cost (cf. Froot et al. (1993), p. 1633). In this paper, we refer to this as uncoordinated or decentralized credit policy.

Firms using bank credit also have a significant leeway in their choice of debt maturity; in a typical U.S. firm, 50% of total borrowing is short-term bank credit and 34% is long-term debt (cf. Rajan and Zingales (1994), p. 1428). One incentive for firms to assume long-term debt is to improve their management of liquidity shocks and short-term borrowing costs (cf. Tirole (2006), Ch. 5). Finally, from the perspective of the firm’s owners, trade and bank credit may have very different outcomes in the event of bankruptcy. With trade credit, the supplier/lender simply repossesses the goods against which credit was granted and the bankruptcy is essentially costless (cf. Peterson and Rajan (1997), p. 662). With bank credit, the bankruptcy is typically costly (for an overview of bankruptcy costs, see Grinblatt and Titman (2002), Ch. 16) and frequently the firm’s owners may have to forfeit their claim to future earnings of the firm (for empirical evidence, see Bernstein (2007)). We refer to this as costly wipeout bankruptcy.

The inventory literature has yet to explore many of these nuances associated with bank credit. The thesis underlying this paper is that they are important because, as a consequence of these nuances, in a market in which firms pay taxes and face the risk of costly bankruptcy, firms ought to adapt their inventory policies in non-trivial ways that are sometimes counterintuitive. For example,
our results reveal that firms that fail to coordinate credit and material flows consistently underinvest in inventory and are not sufficiently leveraged.

We address these issues by using a dynamic model to compare all-equity firms that base their stocking decisions on profit maximization, i.e., well-established inventory theory, with their counterparts that coordinate inventory and financial decisions in order to maximize the market value of the firm's equity. The financial decisions include the amount of a short-term loan and the amount of a dividend to issue net of capital subscriptions. Later we augment the firm’s capital structure with long-term debt. If demand is too low to generate sufficient sales revenue to allow the firm to repay its debts, then the firm declares bankruptcy. We assume that in such a case, the shareholders lose their claim to future earnings of the firm, and the debt holders become the residual claimants. The event of bankruptcy imposes two kinds of extra costs on the operations of the firm: direct costs, such as legal fees and court costs, and indirect costs, such as a loss of revenues if the firm’s operations are inhibited. In our model we do not distinguish between direct and indirect costs of bankruptcy but simply assume that there are bankruptcy costs. In what follows, we place the model in the context of the existing literature and outline our contributions.

1.2. Literature

The classic inventory literature, as described in Graves et al. (3rd reprint 2004) for example, treats inventory decisions independently of financing decisions and can be essentially regarded as a response to Modigliani and Miller (1958). Most papers on inventory financing fall into the class of exceptions to the Modigliani-Miller theorem. Our paper relies on the presence of taxes and costly bankruptcy to depart from the Modigliani-Miller world.

Inventory models come in many varieties, which depend on assumptions about key variables such as demand, costs, and physical aspects of the system. Including financial considerations further increases the number of possible models. In general, the assumptions that one makes are about bankruptcy, and the firm’s objective determines the essential structure of the model. We discuss each of these considerations in turn and then comment on their disposition in the existing literature.

1. Bankruptcy and bankruptcy costs. Modeling bankruptcy is a difficult business and there are several contending approaches to it. As a means of classification, models of bankruptcy can mainly be split into two groups – limited liability at bankruptcy (or ‘wipeout’ bankruptcy) and unlimited liability at bankruptcy (or restructuring bankruptcy). The former, wipeout bankruptcy, assumes that the shareholders lose their claim to future earnings of the firm and the lenders become the residual claimants. This approach is taken here, in Xu and Birge (2004), and in Kouvelis and
Zhao (2009). Restructuring bankruptcy assumes that the shareholders retain their claim to future earnings of the firm, but incur restructuring costs. This approach is taken in Li et al. (1997).

There are two primary ways to model the extra costs that the event of bankruptcy imposes on the operations of the firm. One is to assume that these costs are fixed as is done here and in Brander and Lewis (1988); the other way assumes proportional bankruptcy costs as is done in Li et al. (1997) and Xu and Birge (2004). Both approaches are well-accepted in the economics literature and affect the optimal inventory levels. For reasons that become clear in §4.3, firms that are subject to fixed bankruptcy costs tend to carry more inventory than firms that confront proportional bankruptcy costs.

2. The firm’s objective. In traditional inventory models, firms ordinarily maximize expected profit realized by the firm’s owners. The inclusion of external financing introduces another party with a stake in the firm, namely the lender. One approach is to assume that the firm maximizes the value of the firm’s equity plus the value of its debt. In other words, the criterion is the total expected returns to shareholders and creditors as is used in Xu and Birge (2004). The other approach is to assume that the firm maximizes only the value of equity, as is done here and in Li et al. (1997) and Xu and Birge (2006). Although both approaches are well-accepted in the literature, Brander and Lewis (1986) observe that “...a more standard approach would be to assume that the firm maximizes the value of equity only, since debt holders have no direct control over the management of the firm.” The implication of the firm’s objective on its output is discussed in detail in Brander and Lewis (1986), who observe a higher level of output for firms that maximize the expected value of equity alone.

To the best of our knowledge, the inventory literature with financial constraints has thus far almost exclusively focused on the newsvendor model. The standard reference on the static newsvendor with financial constraints is Xu and Birge (2004). A similar analysis can be found in Kouvelis and Zhao (2009), who consider a simple supply chain with bank credit. Xu and Birge (2004) employs a model with wipeout bankruptcy and proportional bankruptcy costs, and the firm strives to maximize the expected value of equity plus the value of debt. Xu and Birge (2004) establish optimal borrowing and inventory levels, and show that when the firm fails to coordinate its operational and financial decisions, it borrows too little and carries too much inventory. For reasons that become clear in §4.3, the last result is contrary to the conclusions in this paper.

Among the dynamic models, Xu and Birge (2006), Li et al. (1997), and Gupta and Wang (2009) are the closest to ours. Xu and Birge (2006) study optimal inventory-ordering and financial decisions of a firm that maximizes the expected value of equity over a finite operating horizon in the presence
of proportional bankruptcy costs. The model presented in the paper, namely an integer stochastic programming model with non-linear constraints, is not analytically tractable, but the authors develop an efficient algorithm to solve it numerically. They demonstrate numerically that in the presence of bankruptcy costs, coordinated decisions dominate uncoordinated ones.

There are more similarities between the models in the present paper and in Li et al. (1997); we exploit the methods in that paper to answer different questions. The key difference in the models is the treatment of the aftereffects of bankruptcy. The present paper is consistent with limited liability at bankruptcy, whereas the other two papers are consistent with unlimited liability at bankruptcy. The modeling of the aftereffects of bankruptcy profoundly affects the properties of optimal decisions. In a firm that risks reorganization bankruptcy and coordinates its short-run financial and operational decisions, both the inventory level and the amount of the short-term loan should be less than the analogous quantities in a firm that decentralizes its short-run decisions while risking unlimited liability at bankruptcy. In this paper, we find that the inequalities are reversed if both firms risk wipeout bankruptcy. A firm which risks reorganization bankruptcy and coordinates its short-run decisions should be less highly leveraged than its counterpart which also risks reorganization bankruptcy but decentralizes its short-run decisions. In this paper, we find that the inequality would be reversed if both firms risked wipeout bankruptcy.

In a recent paper, Gupta and Wang (2009) characterize an optimal inventory policy of a firm that uses trade credit to finance inventory. Rather than a bank, the source of financing is an infinitely rich supplier. Although the supplier imposes financing charges on the firm, the analysis implicitly assumes that firm always has sufficient funds to cover the charges, and that precludes the possibility of bankruptcy. The paper shows that the structure of an optimal policy is unaffected by these financing charges, but that the inventory level should be adjusted to match the credit terms offered by the supplier. In addition to a credit line or supplier financing, firms may use asset backed financing; this is analyzed in Buzacott and Zhang (2004) who consider a problem of a growing manufacturer who finances production with loans secured by inventory.

1.3. Contribution and Organization of the Paper

We investigate a dynamic newsvendor model with a liquidity constraint and short-term borrowing and, later in the paper, long-term borrowing. We make three contributions in this stylized setting.

First, we characterize an optimal policy for coordinating operational and financial decisions under the risk of costly wipeout bankruptcy. Results under wipeout bankruptcy are useful because apparently in many bankruptcies, the existing shareholders are forced to relinquish their claim to future earnings of the firm (cf. Bernstein (2007)). However, we know of only one-period models
with wipeout bankruptcy thus far. Here we advance the analysis by recognizing that firms do not operate just for one period. It is well-known (cf. Porteus (2002)), that a dynamic newsvendor problem is "...more than $N$ copies of the original newsvendor problem ... the system is operated over $N$ periods ... the leftover stock at the end of one period is retained in inventory and can be offered for sale the following period." Financing and the risk of bankruptcy only complicate the problem further. In particular, with bankruptcy, the newsvendor's expected payoff is not necessarily a concave function of the order quantity. However, we show that the optimal inventory policy is still base-stock.

Second, we compare the optimal coordinated and decentralized policies. As it becomes clear in §4.3, firms that do not coordinate their credit and inventory flows opt for short-term loans that are too small and inventory levels that are too low. We then show that this insight remains valid with non-stationary demand processes.

Third, we investigate the sensitivity of these conclusions to the amount of long-term debt. Since the degree of leverage affects short-term decisions and, therefore, the market value of equity, we ask: what amount of long-term debt maximizes that value? We find that the coordinating firm should be more highly leveraged than the decentralized firm. The reason is that the decentralized firm is loath to exploit the added liquidity that greater leverage would provide.

The paper is organized as follows. In §1.2 we relate this paper to the emerging literature on interactions of operational and financial decisions. The model and preliminary results are in §2. Notable features of the model include random product demand and, in the event of default, liquidation of the firm and cessation of operations. Thus, the structure is an optimization model that is dynamic and stochastic and that has a random stopping time.

In §3, we establish key properties of a dynamic program that corresponds to the optimization problem, and show that there is a myopic optimum (Proposition 1) at which the firm prefers to finance all expenditures with internal funds before turning to external sources (Proposition 2).

In §4, we compare the optimal policies in coordinated and decentralized firms. Inventory decisions in the latter are made to optimize profit, and financial decisions are made by taking as given the cash flows resulting from operations. The optimal policy in the coordinated firm entails larger short-term loans, higher inventory levels, larger dividends, and a higher probability of default than in the decentralized firm (Proposition 3).

The model is generalized in §5 by replacing the sequence of independent and identically distributed random product demands with an additive non-stationary demand process. The principal
negative result is that there is no longer a myopic optimal policy. The principal positive result is that the conclusions in the previous paragraph remain valid.

In §6, we examine the effects of the firm's long-term debt on its short-term financial and operational decisions. We show that a coordinating firm with greater financial leverage should have a higher inventory level and adjust its short-term debt to have a larger default probability than its decentralized counterpart (Proposition 4). Moreover, if the interest cost of a short-term loan is an affine function of the loan size, then the coordinated firm is better off with a greater amount of long-term debt than its decentralized counterpart (Proposition 5).

In §7, we provide economic and behavioral justifications for some of our key assumptions. The contents include a borrower's condition and a lender's condition that imply assumptions that we make in earlier sections. Our conclusions are summarized in §8. Many proofs are relegated to an appendix.

2. The Model

Consider a leveraged corporate firm that strives to maximize its market value. Following standard assumptions in financial economics, this implies that the firm's shareholders strive to maximize the expected value of equity, while the firm's lenders expect to break even. In the event of bankruptcy, the shareholders lose their claim over the earnings of the firm and the lenders become residual claimants.

2.1. Basic Model

We use a discrete-time, multi-period model of a single-product single-location firm which 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. We assume (until §5) that successive demands $D_1$, $D_2$, ... are independent and identically distributed non-negative random variables. To simplify the exposition, we assume that the demand distribution has a density function $f$. However, this assumption is not necessary for the validity of our results. We use $F$ to denote the distribution function.

At the beginning of each period $n$, ($n = 1, 2, \ldots$) the firm observes the amount of retained earnings, $w_n$ (unconstrained in sign), the current inventory level, and $x_n \geq 0$ (implying that excess demand is lost). If $w_{n+1} < 0$, then the firm was unable to meet some of its debt obligations in period $n$ and it is assumed to be in bankruptcy. This means that the shareholders lose their claim over the earnings of the firm and the lenders become residual claimants. Our bankruptcy assumptions are detailed in §2.2.
Prior to observing demand in period \( n \), the firm makes three decisions: a production (or procurement) quantity, the amount of a short-term loan, and the amount of a dividend to declare. These are denoted as \( z_n \geq 0 \), \( b_n \geq 0 \), and \( v_n \) (unconstrained in sign; we interpret negative values of \( v_n \) as capital subscriptions), respectively.

The production and borrowing costs are \( cz_n \) and \( \rho(b_n) \), respectively, where \( c \) denotes the unit production cost, and \( \rho(b_n) \) is the total interest paid on a loan of amount \( b_n \). Often we write \( \rho_n \) for \( \rho(b_n) \) for notational simplicity. Since the probability of bankruptcy increases with the amount borrowed, \( \rho(\cdot) \) is convex and increasing. See §7, for a detailed rationale for this assumption.

We assume that the production lead time is negligible; so

\[
y_n = x_n + z_n
\]

(1)

is the supply of goods that is 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 \( n \). The paper's formulas are based on the assumptions that excess demand is lost and holding costs are linear. If a period's supply of goods is \( y \) and its demand is \( D \), then its sales revenue net of holding costs is

\[
r \min\{y, D\} - h(y-D)^+ = ry - (r+h)(y-D)^+.
\]

(2)

We interpret \( r \) and \( h \) as the respective exogenous sales price and the unit holding cost. When we discount, we assign this gross profit to the end of the period.

Let \( \tau \) be the corporate tax rate. The flow of goods and dollars in the model is subject to conservation constraints

\[
x_{n+1} = (y_n - D_n)^+,
\]

(3)

\[
w_{n+1} = w_n - v_n + (1-\tau)[ry_n - (r+h)(y_n - D_n)^+ - cz_n - \rho_n].
\]

(4)

In §7 we discuss the rationale for the assumption that excess demand is lost rather than backordered.

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

\[
w_n + b_n \geq v_n + (1-\tau)(cz_n + \rho_n).
\]

(5)

The logical constraints are

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

(6)
We employ accrual tax rules to simplify the exposition. That is, a tax is levied if the firm makes a profit, and a tax credit is awarded if it incurs a loss. The accrual rule simplifies the problem while accurately capturing the tax-sheltering advantage of debt financing, a reason to include corporate taxes in the model.

By the end of period \( n \), the firm will have observed demand \( D_n \), realized revenue net of inventory costs, and repaid the entire loan principal \( b_n \) if \( w_{n+1} \geq 0 \), where \( w_{n+1} \) is given by (4). Otherwise, the firm must declare its inability to pay its debts and it enters into bankruptcy.

**Replacing Flow Variables with Level Variables.** As in Li et al. (1997), we reduce the dimensionality of the state vector by replacing the flow variables \( z_n \) and \( u_n \) with new variables that specify process levels after the period \( n \) decisions are implemented. The first of those level variables is \( y_n \) defined in (1), and the second is

\[
s_n = w_n - v_n - (1 - \tau)(cz_n + \rho_n) + b_n. \tag{7}
\]

The production quantity, \( z_n \), is replaced by the supply level \( y_n \) and the dividend, \( u_n \), is replaced by \( s_n \). The latter is the working capital after paying the dividend, loan interest, and production costs, but before inventory costs are realized.

This replacement yields a simple form of the liquidity and logical constraints (5) and (6):

\[
s_n \geq 0, \quad b_n \geq 0, \quad \text{and} \quad y_n \geq x_n. \tag{8}
\]

### 2.2. Bankruptcy and Bankruptcy Costs

In practice, the meaning of bankruptcy is that a firm enters into a legal process as a result of declaring its inability to pay its debts as they come due. Some firms can continue to operate while in bankruptcy, and others are liquidated. In either case, the shareholders lose their claim to the immediate earnings of the firm, and the lenders become residual claimants.

The advent of bankruptcy imposes extra costs on the operations of the firm, and we assume that the firm incurs a fixed bankruptcy cost, \( \mathcal{L} \). See Grinblatt and Titman (2002), Ch. 16, for an overview of bankruptcy costs. Who bears the bankruptcy cost, \( \mathcal{L} \)? Under the *absolute priority rule*, most of the firm’s value in the event of bankruptcy is transferred to its lenders. Since the cost of bankruptcy diminishes the value of the firm, most bankruptcy costs are ultimately borne by the firm’s lenders. However, knowing that they would have to bear a potential loss, lenders adjust the interest rate and, therefore, increase the equity holders’ borrowing costs. *The bankruptcy risk implications for the total borrowing cost are discussed in §7.1* (the total borrowing cost is convex and increasing in the loan size). The scope of the current section is restricted to the cash flows.
From the definition of \( s_n \) in (7), bankruptcy occurs at the end of period \( T \) where

\[
T = \inf \{ t : s_t + (1 - \tau)[r y_t - r(y_t - D_t)^+] \leq b_t \}. \tag{9}
\]

That is, \( T \) is the first time that \( w_{T+1} \leq 0 \). In words, bankruptcy in the model occurs when the period-\( n \) revenue is not sufficient to cover promised payments to the lenders. This bankruptcy definition is consistent with the convention adopted in Xu and Birge (2004). (Alternatively, we could have assumed that in addition to the sales revenue, the firm can also use the salvage value of any unsold inventory to cover outstanding debt payments. This assumption, however, would not affect any qualitative insights given throughout the paper.)

If \( y_n = y, b_n = b \) and \( s_n = s \), let \( q(y, b, s) \) be the conditional probability that bankruptcy does not occur in period \( n \), given that it has not occurred prior to period \( n \):

\[
q(y, b, s) = P \{ s + (1 - \tau)[r y - r(y - D)^+] \geq b \} = P \{ (1 - \tau)[r y - r(y - D)^+] \geq b - s \}, \tag{10}
\]

where \( P \) denotes probability. To avoid uninteresting cases, we assume that \( y \) and \( b \) are chosen so that bankruptcy does not occur if all goods are sold, i.e. if \( D \geq y \). This implies

\[
(1 - \tau)r y > b - s. \tag{11}
\]

Then the dependence on \( y \) vanishes and (10) becomes

\[
q(b, s) = P \left\{ D > \frac{b - s}{r(1 - \tau)} \right\}. \tag{12}
\]

Now, define

\[
m(b, s) = \frac{b - s}{r(1 - \tau)}, \tag{13}
\]

and notice that (11) and (13) imply \( y > m \). Bankruptcy can occur only if \( D < m \), so bankruptcy occurs with certainty if \( y \leq m \). Hence, we assume that \( y > m \).

In bankruptcy, after liquidating the inventory and deducting the bankruptcy costs, the total amount available to claimants is:

\[
W = (1 - \tau)[- (r - c_S)(y_T - D_T)^+ + r y_T - \mathcal{L}] + s_T = (1 - \tau)[c_S(y_T - D_T)^+ + r D_T - \mathcal{L}] + s_T, \tag{14}
\]

where \( c_S \) is the fire-sale price of residual inventory and is often much smaller than \( c \).

If the residual monies are sufficiently large, the bank receives up to \( b_T + \rho_T \). That is, the lenders receive:

\[
\mathcal{I} = \min \{ b_T, W^+ \}. \tag{15}
\]

The shareholders receive the residual:

\[
\mathcal{J} = (W - \mathcal{I})^+ = (W - b_T - \rho_T)^+ = \{ (1 - \tau)[c_S(y_T - D_T)^+ + r D_T - \mathcal{L}] + s_T - b_T \}. \tag{16}
\]
2.3. Optimization Objective

Let the scalar $\beta$ denote an exogenously determined single-period discount factor, $0 < \beta < 1$. The determination of an appropriate $\beta$ is affected by the presence of corporate income taxes; see §13 in Grinblatt and Titman (2002). We omit these details because they do not affect the qualitative nature of our results. The present value of the dividends, net of capital subscriptions, is

$$V = \sum_{n=1}^{T} \beta^{n-1}v_n + \beta^{T-1}J.$$  \tag{17}

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 assigns feasible values to $z_n$, $b_n$, and $v_n$ for each $n$, i.e., it is a function that assigns a feasible value to $(z_n, b_n, v_n)$ for each possible $H_n$ and $n$. A policy is optimal with respect to a set of initial states $S$ if it maximizes $E(V|H_n)$ for each $H_n$ and $n$ such that $(x_1, w_1) \in S$. In the remainder of the paper we characterize a policy that is optimal with respect to an appropriate $S$. First, we obtain a compact expression for $E(V)/(1-\tau)$.

Equations (4), (7), and the identities $z_n = y_n - x_n$, $x_n = (y_{n-1} - D_{n-1})^+$ (if $n > 1$) yield

$$v_n = (1-\tau)[ry_{n-1} - (r + h)(y_{n-1} - D_{n-1})^+ - \rho_n - cy_n + c(y_{n-1} - D_{n-1})^+] + b_n - b_{n-1} - s_n + s_{n-1}$$

$$= (1-\tau)[ry_{n-1} - (r + h - c)(y_{n-1} - D_{n-1})^+ - \rho_n - cy_n] + b_n - b_{n-1} - s_n + s_{n-1}.$$  \tag{18}

Therefore,

$$V = \sum_{n=1}^{T} \beta^{n-1}v_n + \beta^{T-1}J = w_1 - (1-\tau)(\rho_1 + cy_1 - cx_1) + b_1 - s_1$$

$$+ \sum_{n=2}^{T} \beta^{n-1}\{(1-\tau)[ry_{n-1} - (r + h - c)(y_{n-1} - D_{n-1})^+ - \rho_n - cy_n] + b_n - b_{n-1} - s_n + s_{n-1}\}$$

$$+ \beta^{T-1}\{(1-\tau)[cS(y_T - D_T)^+ + ry_T - r(y_T - D_T)^+ - L] + s_T - b_T\}^+.$$  

Rearranging terms yields

$$V = w_1 + (1-\tau)cx_1 + \sum_{n=1}^{T-1} \beta^{n-1}\{(b_n - s_n)(1-\beta) + (1-\tau)[(\beta r - c)y_n - \beta(r + h - c)(y_n - D_n)^+ - \rho_n]\}$$

$$+ \beta^{T-1}[b_T - s_T - (1-\tau)(\rho_T + cy_T)] + \beta^{T-1}\{(1-\tau)[rD_T - cS(y_T - D_T)^+ - L] + b_T + s_T\}^+.$$  

Hence,

$$\frac{V}{1-\tau} = \frac{w_1}{1-\tau} + cx_1 + \sum_{n=1}^{T-1} \beta^{n-1}\alpha_1(y_n, b_n, s_n, D_n) + \beta^{T-1}\alpha_2(y_T, b_T, s_T, D_T)$$  \tag{19}

with the definitions

$$\alpha_1(y, b, s, u) = \left(\frac{1-\beta}{1-\tau}\right)(b - s) - \rho(b) + (\beta r - c)y - \beta(r + h - c)(y - u)^+,\tag{20a}$$
\[
\alpha_2(y, b, s, u) = \frac{b-s}{1-\tau} - \rho(b) - cy + \left[ ru + c_s(y-u)^+ - \mathcal{L} - \frac{b-s}{1-\tau} \right]^+.
\] (20b)

As discussed in §2.2, default occurs if and only if \( u < m \), so \( \tau \geq c_s \) implies that an upper bound on \([\ldots]^+\) in (20b) is

\[
\left[ ru + c_s(y-u)^+ - \mathcal{L} - \frac{b-s}{1-\tau} \right]^+ \leq \left[ rm + c_s(y-m) - \mathcal{L} - \frac{b-s}{1-\tau} \right]^+.
\]

Hence, if

\[
\mathcal{L} \geq c_s(y-m) + rm
\] (21)

then \([\ldots]^+\) in (20b) is zero. Namely, (21) implies that the shareholders receive nothing after the firm is liquidated. In the subsequent analysis we assume that (21) is valid, and we simplify (20b) with

\[
\alpha_2(y, b, s, u) = \frac{b-s}{1-\tau} - \rho(b) - cy.
\] (22)

This assumption streamlines the exposition, but it does not affect the validity of any of our results. Empirical evidence that supports this assumption can be found in Bernstein (2007). As a consequence of (19), (20a), and (22),

\[
\frac{E(V)}{1-\tau} - \frac{w_1}{1-\tau} - cx_1 = E \left[ \sum_{n=1}^{T-1} \beta^{n-1} \alpha_1(y_n, b_n, s_n, D_n) + \beta^{T-1} \alpha_2(b_T, y_T, D_T) \right] = \sum_{n=1}^{\infty} \beta^{n-1} P\{T > n-1\} \left( E[\alpha_1(y_n, b_n, s_n, D_n)|T > n, T > n-1]P\{T > n|T > n-1\} \right.
\]

\[
+ E[\alpha_2(b_n, y_n, D_n)|T = n, T > n-1]P\{T = n|T > n-1\} \right).
\]

Since \( \{u : (1-\tau)(ry - (r+\hat{h})(y-u)^+) > b-s\} = (m, \infty) \), (11) implies

\[
E[\alpha_1(y_n, b_n, s_n, D_n)|T > n]P\{T > n|T > n-1\}
\]

\[
+ E[\alpha_2(y_n, b_n, s_n, D_n)|T = n, T > n-1]P\{T = n|T > n-1\}
\]

\[
= E_{b_n,y_n} \left[ q(b_n, s_n) \int_{m(b_n,s_n)}^{\infty} \alpha_1(y_n, b_n, s_n, u)dF(u) \right. \left. + [1-q(b_n, s_n)] \int_{0}^{m(b_n,s_n)} \alpha_2(y_n, b_n, s_n, u)dF(u) \right].
\]

Therefore, we have the following compact expression for \( E(V)/(1-\tau) \):

\[
\frac{E(V)}{1-\tau} - \frac{w_1}{1-\tau} - cx_1 = E \left( \sum_{n=1}^{\infty} \beta^{n-1} \left[ \prod_{m=1}^{n-1} q(b_{m}, s_{m}) \right] K(y_n, b_n, s_n) \right),
\] (23)

where

\[
K(y, b, s) = q(b, s) \int_{m}^{\infty} \alpha_1(y, b, s, u)dF(u) + [1-q(b, s)] \int_{0}^{m} \alpha_2(y, b, s, u)dF(u).
\] (24)

It is important that \( K \) does not depend on \( w \) and \( x \).
This model can be regarded as a generalization of an inventory process with a stopping time. Therefore, Lovejoy (1992) provides a bound on the error that would result from using a myopic optimal policy with an infinite-horizon. However, we avoid the need for an approximation by showing in the next section that the model with wipeout bankruptcy has an optimal myopic solution.

3. Preliminary Results

It follows from (8) and (19) that an optimal policy does not depend on \( \{w_n\} \). That is, maximizing \( E(\mathcal{V}) \) corresponds to the following dynamic program that has a scalar state variable, \( x \), rather than the vector \( (x, w) \) where \( x \) and \( w \) are the respective levels of inventory and working capital at the beginning of a period:

\[
\phi(x) = \sup\{J(y, b, s) : s \geq 0, b \geq 0, y \geq x\}, \quad (25a)
\]

\[
J(y, b, s) = K(y, b, s) + \beta q(b, s) \int_{m}^{\infty} \phi([y - u]^{+}) dF(u). \quad (25b)
\]

Notice that \( x \) plays no role on the right side of (25b). So this MDP satisfies the conditions in Sobel (1981) and, if it has an optimal policy, then it has a myopic optimal policy. We now explore the structure of this policy. If action \( (y_n, b_n, s_n) = (y, b, s) \) is feasible and taken in all periods \( n \), then

\[
\frac{E(\mathcal{V})}{1 - \tau} - \frac{w_1}{1 - \tau} - cx_1 = K(y, b, s)[1 + \beta q(b, s) + \beta^2 q^2(b, s) + \cdots] = K(y, b, s)/[1 - \beta q(b, s)].
\]

Let \( (y, b, s) = (y_*, b_*, s_*) \) denote a triple at which \( K(y, b, s)/[1 - \beta q(y, b, s)] \) achieves its global maximum subject to \( b \geq 0 \) and \( s \geq 0 \) (which we assume exists). Then

\[
\phi(x) = K(y_*, b_*, s_*)/[1 - \beta q(b_*, s_*)] \quad x \leq y_*.
\]

This leads to the following result (in which the policy stipulates an arbitrary feasible action in any period \( n \) where \( x_n > y_* \)).

PROPOSITION 1. If \( K(y, b, s)/[1 - \beta q(b, s)] \) achieves its global maximum (constrained by \( b \geq 0 \) and \( s \geq 0 \)) at \( (y, b, s) = (y_*, b_*, s_*) \), then \( (y_n, b_n, s_n) = (y_*, b_*, s_*) \), \( n = 1, \cdots, T \), is an optimal policy with respect to the set of initial states \( S = (-\infty, y_*) \). Under this policy, the lifetime of the firm \( T \) has a geometric distribution with parameter \( q(b_*, s_*) \).

That is, if the initial inventory level is not too high \( x_1 \leq y_* \), then it is feasible and optimal for \( (y_n, b_n, s_n) = (y_*, b_*, s_*) \) for all \( n \). This exposes the firm each period to the probability \( 1 - q(b_*, s_*) \) of bankruptcy. So the resulting lifetime of the firm \( T \) is has a geometric distribution with parameter \( q(b_*, s_*) \).
This result makes clear that the firm should finance all expenditures with internal funds before turning to external sources. Consistent with Li et al. (1997), the firm should borrow the smallest amount that satisfies the liquidity constraint (8).

**Proposition 2.** In (25), \( s_n b_n = 0, \ n = 1, \ldots, T \) is optimal.

Two cases are consistent with the constraints (8) and this complimentary slackness condition. One is \( s_n \geq 0 \) and \( b_n = 0 \) which is feasible but uninteresting because, due to Proposition 2, it reduces to a standard dynamic newsvendor problem. The other is \( s_n = 0 \) and \( b_n \geq 0 \) *which we assume henceforth.*

With \( s = 0 \), it is apparent from (13) that \( m \) and \( b \) determine each other. It is convenient now to let \( m \) and \( y \) be the decision variables and treat \( b \) as a dependent variable given by

\[
b = (1 - \tau)rm. \tag{27}\]

Henceforth, we abuse notation and write \( q(m) \) instead of \( q[(1 - \tau)rm, 0] \).

Using specification (24) of \( K \), let

\[
K(y, m) = K[y, (1 - \tau)rm, 0] = [1 - F(m)] \int_{m}^{\infty} \alpha_1(y, (1 - \tau)rm, 0, u) dF(u) + F(m) \int_{0}^{m} \alpha_2(y, (1 - \tau)rm, 0, u) dF(u). \tag{28}\]

The definitions of \( \alpha_1 \) and \( \alpha_2 \) ((20a) and (22)) and the assumption \( \mathcal{L} \geq c_s(y - m) + rm \) yield

\[
\alpha_1[y, (1 - \tau)rm, 0, u] = (1 - \beta)rm - \rho[(1 - \tau)rm] + (\beta r - c)y - \beta(r + h - c)(y - u)^+, \tag{29a}\]

\[
\alpha_2[y, (1 - \tau)rm, 0, u] = rm - cy - \rho[(1 - \tau)rm]. \tag{29b}\]

Finally, it is convenient to define \( m_* = b_*/[\tau(1 - \tau)] \) and

\[
L(y, m) = \frac{K(y, m)}{1 - \beta q(m)} = \frac{K[y, r(1 - \tau)m, 0]}{1 - \beta + \beta F(m)} \quad m \geq 0, \ y \in \mathbb{R}. \tag{30}\]

From (26),

\[
\phi(x) = L(y_*, m_*) \quad if \quad x \leq y_*.
\]

In the next section, we compare \( (y_*, m_*) \) with the analogous decisions in decentralized firms.
4. Centralized vs. Decentralized Decision Making

The Modigliani-Miller theorem has become the modern basis of thinking about inventory financing. If the theorem holds, then how the firm finances inventory becomes essentially irrelevant. Therefore, most papers on inventory financing rest on exceptions to the Modigliani-Miller theorem. Our paper relies on the presence of taxes and costly bankruptcy to depart from the Modigliani-Miller world. These assumptions imply that inventory and financing decisions are interdependent. However, since we consider a dynamic model, most of the interactions between debt and inventory would occur even if bankruptcy costs and taxes were zero (for a further discussion, see Ch. 3 in Tirole (2006)).

Since operational and financial decisions are linked then in the absence of coordination costs, a centralized policy that coordinates inventory and financial flows ought to dominate any uncoordinated policy. In practice, the coordination costs may not be zero, but thanks to advances in information technology they have become sufficiently low for coordination to become economically interesting. Decentralized operational and financial decisions, however, are the modus operandi at many large firms. Casual observations in practice (e.g. inventory management at Emerson Electric) suggest that operations managers do not consider financing. First, they formulate an inventory decision, and then they regard the job of financial managers as that of securing the necessary financing to accommodate it. Therefore, finance coordinates with operations, but not vice versa; Li et al. (1997) refers to this phenomenon as “one-sided coordination.”

Our main result in this section reveals that the firms that engage in “one-sided coordination” do not borrow enough, and they consistently underinvest in inventory. The intuition behind this result is that the decentralized policy fails to consider the benefits of leverage and limited liability. The section proceeds with definitions of centralized and decentralized decision making, and then compares the optimal policies in both cases. In the subsequent sections, we show that the result remains valid when the stochastic demand becomes non-stationary (e.g. due to seasonality) and the firm uses several classes of debt (namely, short- and long-term debt).

4.1. Decentralized Operational and Financial Decisions

We use the superscript “U” for uncoordinated (or decentralized) and assign an expected present value of dividends (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 Modigliani-Miller 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, namely the expected present value of revenues net of inventory- and production-related costs:
\[
\sum_{n=1}^{\infty} \beta^{n-1} \{\beta [ry_n - (r + h)(y_n - D_n)^+] - cz_n\} = \sum_{n=1}^{\infty} \beta^{n-1} [\beta(r - c)y_n - \beta(r + h - c)(y_n - D_n)^+] + cz_1 - \beta c \sum_{n=1}^{\infty} \beta^{n-1} D_{n+1}.
\]

Thus, the operational subproblem in the decentralized case is the non-anticipative choice of \(y_1, y_2, \ldots\) to
\[
\max E \left\{ \sum_{n=1}^{\infty} \beta^{n-1} \{\beta [ry_n - (r + h)(y_n - D_n)^+] - cz_n\} \right\} \quad \text{(subject to } y_n \geq x_n \text{ for all } n). \tag{31}
\]

This familiar dynamic newsvendor model has the following well-known myopic optimum (cf. Heyman and Sobel (2004a)). Choose \(y = y^U\) to maximize \((\beta r - c)y - \beta(r + h - c)E[(y - D)^+]\). Therefore, \(y^U\) satisfies
\[
F(y^U) = \frac{\beta r - c}{\beta(r + h - c)}, \tag{32}
\]
where \(F\) denotes the distribution function of \(D\). If \(x_1 \leq y^U\), i.e., if the initial inventory is not too high, then \(y_n = y^U\) for all \(n\) is feasible and optimal in the operational subproblem \(U\).

The financial subproblem in the decentralized case treats as given the cash flows which result from operational decisions, and non-anticipatively chooses the short-term loans and dividends to maximize the EPV of dividends net of bankruptcy proceeds. So this subproblem uses \(y_n = y^U\) for all \(n\) and it corresponds (due to Proposition 1) to choosing \(m \geq 0\) to maximize \(L(y^U, m)\), given by (30). We constrain \(m \geq 0\) because \(m = b/[(1 - \tau)r]\) with \(b \geq 0\). Let \(m^U\) maximize \(L(y^U, \cdot)\) on \([0, \infty)\), and let \(b^U = (1 - \tau)rm^U\).

### 4.2. Centralized Operational and Financial Decisions

We use the superscript "C" for coordinated (or centralized). The coordinated problem consists of \(\textit{jointly}\) selecting \(\{y_n\}\) and \(\{m_n\}\) to maximize the EPV of dividends. Accordingly, let \((y, m) = (y^C, m^C)\) achieve
\[
\max \{L(y, m) : 0 \leq y, 0 \leq m \leq m^*\}. \tag{33}
\]

It follows from Proposition 1 that \((y_n, b_n) = [y^C, (1 - \tau)rm^C], n = 1, 2, \ldots,\) with \(x_1 \leq y^C,\) is a myopic optimal policy in problem \(C\).

\textbf{Notation:} If \(G(\cdot, \cdot)\) is twice differentiable, we write \(G^{(1)}(y, m)\) and \(G^{(2)}(y, m)\) to denote the left-hand partial derivative of \(G(y, m)\) with respect to \(y\) and \(m\), respectively, and \(G^{(12)}(y, m)\) to denote the cross-partial derivative with respect to \(y\) and \(m\).
4.3. Comparison of Coordinated and Uncoordinated Solutions

The purpose of this section is to compare optimal policies in the decentralized and centralized cases.

For a fair comparison, we conduct the analysis on a reasonable set of parameters, which includes imposing a capacity on borrowing. Intuitively, the firm’s lenders expect to earn an acceptable rate of return on the short-term loan. However, as the loan amount surpasses some critical amount, no rate of interest charged to the borrower is sufficient to ensure this acceptable expected rate of return. This critical amount must, therefore, be the firm’s borrowing capacity. We formalize this idea in §7, where we give the rationale for the constraint $m \leq m^*$, where

$$m^* = F^{-1} \left[ \frac{\beta h + (1 - \beta)c}{2[\beta h + (2 - \beta)c]} \right].$$

(34)

In terms of the original decision variables, the constraint corresponds to $b = (1 - \tau)rm \leq (1 - \tau)rm^*$. The general idea is that creditors would shun the firm if $m > m^*$ because they would find the risk of default unacceptably high. Similarly, the firm would decline to borrow due to unreasonably high borrowing costs.

A firm enters bankruptcy in period $n$ if it has survived periods $1, 2, \ldots, n - 1$ and if $D_n \leq m = b/[(1 - \tau)\tau]$. So the (conditional) default probability is $F(m)$. Imposing the constraint $m \leq m^*$, the default probability is bounded above by

$$F(m^*) = \frac{\beta h + (1 - \beta)c}{2[\beta h + (2 - \beta)c]}.$$

The next result shows that coordination results in higher levels of borrowing and inventory, and it sheds some light on the link between financing and inventory.

**Proposition 3.** (i) $y^C \geq y^U$, and (ii) $b^C \geq b^U$.

**Discussion.** We interpret $(y^U, m^U)$ and $(y^C, m^C)$ as an idealization of decisions in a well-managed functionally decentralized and functionally centralized firm, respectively. Part (ii) of the Proposition 3 essentially confirms results found elsewhere in the literature: namely that a centralized policy increases the firm’s borrowing capacity. This is because a centralized policy allows the firm to find an optimal tradeoff between the benefits of debt and distress costs incurred in the event of bankruptcy (for example, see Xu and Birge (2004)).

Discussion that follows Corollary 1, discusses the less intuitive part of Proposition 3, namely part (i). Corollary 1 makes it explicit that borrowing and inventory are indeed interdependent and inventory is an increasing function of borrowing.

**Proof of Proposition 3.** The proof uses the fact that $L(\cdot, m)$ has a left-hand partial derivative, $L(\cdot, m)$, and that is a consequence of the following result which is proven in the Appendix.
Lemma 1. \( L(\cdot, m) \) is concave on \( \mathbb{R} \) (for each \( m \geq 0 \)).

(i) We prove \( y^U \leq y^C \) by using the lemma with the inequalities

\[
L^{(1)}(y^C, m^C) \leq 0 \leq L^{(1)}(y^U, m^C).
\] (35)

The left inequality follows from the optimality of \((y^C, m^C)\) in (33), and the right inequality is a consequence of \( m^C \leq m^* \) (because the optimization in (33) has the constraint \( 0 \leq m \leq m^* \)) and \( 0 \leq L^{(1)}(y^U, m) \) if \( 0 \leq m \leq m^* \) as we now show.

From definition (30) of \( L \),

\[
[1 - \beta + \beta F(m)] L^{(1)}(y, m) = K^{(1)}(y, m).
\] (36)

So \( L^{(1)}(y^U, m) \geq 0 \) if \( K^{(1)}(y^U, m) \geq 0 \). Use \( F(y^U) = (\beta r - c)/[\beta (r + h - c)] \) (from (32)) in expression (72) (in the Appendix) for \( K^{(1)} \) to obtain

\[
K^{(1)}(y^U, m) = F(m)\{F(m)[\beta h + c(1 - \beta)] - cF(m)\} \\
= F(m)\{\beta h + c(1 - \beta) - F(m)[\beta h + c(2 - \beta)]\}
\] (37)

whose non-negativity corresponds to \( m \leq m^* \).

(ii) The proof of \( b^C \geq b^U \) uses a supermodularity property which we prove in the Appendix.

Lemma 2. \( L(\cdot, \cdot) \) is a continuous supermodular function on \([y^U, \infty) \times [0, m^*]\).

Supermodularity and continuity of \( L \) and compactness of \([0, m^*]\) imply that there is a selection of \( \hat{m}(y) \in \arg\max\{L(y, m) : 0 \leq m \leq m^*\} \) with \( \hat{m}(\cdot) \) non-decreasing on \([y^U, \infty)\). With this notation, \( m^C = \hat{m}(y^C) \) and \( m^U = \hat{m}(y^U) \). So \( y^U \leq y^C \) from part (i) implies \( m^U \leq m^C \), and

\[
b^U = (1 - \tau) rm^U \leq (1 - \tau) rm^C = b^C. \quad \square
\]

Under mild regularity conditions,

\[
\max\{L(y, m) : y^U \leq y\}
\]

is achieved, say by \( \tilde{y}(m) \), for each \( m \in [0, m^*] \). The argument at the end of the proof yields the following corollary.

Corollary 1. \( \tilde{y}(\cdot) \) is non-decreasing on \([0, m^*]\).
Equivalently, greater financial leverage induces higher inventory levels.

Discussion. The result in Corollary 1 is counterintuitive at first. A firm becomes increasingly vulnerable to default by assuming more debt, and as a consequence of Lemma 3 (cf. Appendix B) its marginal production costs increase. Therefore, should it not offset the heightened risk by reducing its supply level? It is worth mentioning that Corollary 1 is not a consequence of any special assumptions designed to produce counterintuitive results. The intuition behind the result in Corollary 1 is as follows.

The paragraph that follows Proposition 3 argues that when the firm coordinates borrowing and inventory decisions, it tends to borrow more. Due to supermodularity of the myopic objective function \( L \) (cf. (30) and Lemma 2), having more debt increases the marginal returns to having more inventory, and that indicates that the inventory level should be increased. In other words, if the firm has a high level of short-term debt, it is going into bankruptcy unless demand is very high also.

Since the equity holders have limited liability, they have a strategic incentive for the use of debt. If the firm goes into bankruptcy, it does not matter to them whether it misses its debt target by a little or a lot; their claim on the firm effectively goes to zero, the lenders become the residual claimants on the firm’s profits, and the bankruptcy costs incurred in such a case are a fixed amount \( L \). The only relevant consideration for the equity holders is that in the good states of the world (i.e. states with sufficiently high demand), when the firm is only on the verge of bankruptcy (instead of over the edge), its best chance to avoid bankruptcy is to capitalize on favorable demand by carrying sufficient inventory. Therefore it is the additive effect of fixed bankruptcy costs and limited liability at bankruptcy that drives this result.

In a one-period setting with proportional bankruptcy costs and the objective of maximizing the value of the firm’s equity plus the value of its debt, the results in Xu and Birge (2004) imply that \( \tilde{y}(\cdot) \) becomes non-increasing. However, unlike the model considered in this paper, the magnitude of total bankruptcy costs incurred by the lenders at bankruptcy is non-decreasing in the firm’s borrowing and inventory decisions. Since the objective is to maximize the combined cash flows to equity holders and to lenders, these bankruptcy costs are directly inserted into the objective function, marginal costs rise fast with debt level, and the firm finds it optimal to reduce its inventory level as its borrowing increases.

The effect of unlimited liability at bankruptcy combined with proportional bankruptcy costs is illustrated in Li et al. (1997). Similar to Xu and Birge (2004), the results in that paper also imply that \( \tilde{y}(\cdot) \) becomes non-increasing. The model considered in Li et al. (1997) is dynamic with
bankruptcy costs that are increasing with the firm's cash shortfall observed at bankruptcy, and the firm's objective is to maximize only the value of equity. The twist in Li et al. (1997) is that at bankruptcy, the equity holders together with the lenders agree on a debt restructuring plan that requires that the equity holders (not the lenders as is the case here and in Xu and Birge (2004)) pay the bankruptcy costs plus all unpaid debts. In a repeated bankruptcy, these costs are cumulative. In exchange for absorbing the bankruptcy costs, the equity holders retain their claim over the earnings of the firm. Increasing the amount of short-term debt increases both the probability of incurring the bankruptcy costs and their magnitude. Consequently, as short-term borrowing increases, an optimal production level decreases due to an increase in marginal production costs.

5. Non-Stationary Demand

Models of all types impose severe restrictions and suppress myriad realistic details in order to highlight the effects of relatively few parameters. The model in this paper is not an exception, and §7 discusses the basis for some of our key assumptions. However, this section and §6 discuss the effects of relaxing two apparently key assumptions, namely stationarity of demand and a capital structure that is free of long-term debt.

Each result thus far seems to exploit the assumption that the quantities of product demanded in successive periods, \( D_1, D_2, \ldots \), are independent and identically distributed non-negative random variables. Although this assumption pervades the research literature, here we evaluate the robustness of the results in §4 when the sequence of demands is non-stationary.

In general, the probability distribution of each period's demand could depend on decisions and realized demands in previous periods, and on exogenous economic parameters. Two obvious but important examples are demand models with seasonal effects and trends. In order to accommodate the general case, the state of a dynamic program (corresponding to (25)) would be a vector which would include, at least, the entire past history of demands and decisions. It is unlikely that such a dynamic program would be amenable to analysis, and it would not yield a simply structured stocking and borrowing policy.

Here, we consider an additive demand model that decomposes demand into two components: a deterministic component, say \( \nu \), that is a function of demand history, and a stochastic component that captures inherent demand uncertainty. The statistical literature is dominated by time series models that assume additive shocks. Additive models are well understood, relatively easily estimated, encompass many forms of systematic non-stationarity. See Lovejoy (1992) for additive demand processes in operations.

The detailed results are presented in Appendix A and have the following features.
The optimal inventory policy in the decentralized (uncoordinated) setting is determined by a sequence of base-stock levels. This is known under certain conditions due to Veinott (1965) whose non-stationary demand process is more general than additive processes.

In the decentralized regime, let $y^U(\nu)$ be the optimal base-stock inventory level and let $b^U(\nu)$ be the optimal short-term loan. In the centralized regime, let $y^C_t(\nu)$ and $b^C_t(\nu)$ be the corresponding quantities in period $t$. Then Proposition 3 remains valid in the sense that $y^U(\nu) \leq y^C_t(\nu)$ and $b^U(\nu) \leq b^C_t(\nu)$ for all $t$ and $\nu$.

An optimal decentralized policy for short-term decisions is myopic, but an optimal coordinated policy is not myopic. The absence of myopia here is surprising because the "pure" inventory problem with stationary demand does have a myopic optimum (Veinott (1965)).

The optimal coordinated inventory and short-term borrowing policy is determined by base stock levels that linearly increase with the deterministic portion of demand, $\nu$.

In both the decentralized and coordinated cases, in every period $n$, the firm "wants" to satisfy the entire deterministic portion of demand and to engage in additional borrowing in the amount of $(1 - \tau)r\nu$. That amount is the after-tax revenue that stems from the deterministic part of demand. That is, the firm always "wants" to borrow against the "sure" income stream and pass it through to the equity holders as dividends.

6. Long-Term Decisions on Capital Structure

Corporate capital structure is the subject of a large and constantly growing literature. See Allen and Winton (1995), Maksimovic (1995), and Frydenberg (2004) for recent surveys and perspectives on the literature. However, no attention seems to have been given to the effects of capital structure on short-term decisions and, via the effects on those decisions, on the value of the firm. In this section we make an initial foray in that direction.

Thus far, we have analyzed a model of short-term operational and financial decisions in an all-equity firm. Here, we consider the same decisions when the firm's capital structure has long-term debt as well as equity. The first goal is to determine the extent to which the previous insights remain valid when the capital structure includes long-term debt. The second goal is to determine the qualitative impact of the amount of long-term debt on the optimal short-term decisions. The third goal is to compare the amounts of long-term debt that maximize the the market value of the firm in coordinated and decentralized settings. The first step, taken in §6.1, is to generalize the model in §2 to include long-term debt.

We enrich the capital structure in the model with long-term debt having an infinite maturity date. This maturity-date assumption is analytically convenient, occurs previously in theory, and is
approximated in practice. Merton (1974) and Black and Cox (1974) investigate the effects of infinite maturity debt in a dynamic model setting. In practice, not long ago, some firms issued 50-year debt. Two firms (Disney and IBM) issued 100-year debt (cf. Leland (1994), p. 1215). Moreover, if debt has a sufficiently long but finite maturity, the principal can be effectively ignored because it has only negligible value. For example, the repayment of the principal of a 30-year loan with a 10% coupon represents only 5% of the present value of the payments.

6.1. Inserting Long-Term Debt in the Model

Let $Q$ be the amount of a coupon that is due at the end of each period, and let $\mu(Q)$ denote the consequent amount of long-term debt. That is, the firm receives $\mu(Q)$ in return for paying $\$Q$ each period (until bankruptcy occurs). We assume that $\mu(\cdot)$ is concave and increasing; concavity reflects the capital market’s perception of a rising risk of default as the coupon payment grows (equivalently, as the principal of the loan, $\mu(Q)$, grows). Let $\eta$ denote the dollar amount of initial equity, so the firm’s initial capital position is $w_1 = \eta + \mu(Q) \geq 0$.

As in the previous sections, we assume that the firm makes periodic short-term decisions on borrowing, dividends, and production, that those decisions must honor the liquidity constraint (8), and that the chronology is the same with the added stipulation that the coupon payment $Q$ is due at the end of each period. The analogue of (11), which assures that bankruptcy does not occur if all goods are sold, is

$$(1 - \tau)(ry_n - Q) > b_n - s_n. \quad (38)$$

The corresponding replacement of definition (7) is

$$s_n = w_n - v_n - (1 - \tau)(\rho_n + cz_n) + b_n. \quad (39)$$

This yields the following modification of the working capital flow balance equation (4):

$$w_{n+1} = s_n + (1 - \tau)[ry_n - (r + h)(y_n - D_n)^+ - Q] - b_n. \quad (40)$$

Corresponding to definition (9), bankruptcy occurs in period $T$ if that is the first time that the firm’s end-of-period cash does not exceed the short-term loan principal and the post-tax coupon. Specifically, bankruptcy occurs in the earliest period $T$ when

$$s_T + (1 - \tau)r \min\{y_T, D_T\} \leq b_T + (1 - \tau)Q.$$ 

Since bankruptcy does not occur if all of the goods are sold, i.e., if $D_n \geq y_n$ (due to assumption (38)),

$$T = \inf\{n : s_n + (1 - \tau)[ry_n - (r + h)(y_n - D_n)^+] \leq b_n + (1 - \tau)Q\} = \inf\{n : (1 - \tau)rD_n \leq b_n + (1 - \tau)Q\}$$
\[ = \inf \left\{ n : D_n \leq \frac{b_n - s_n + (1 - \tau)Q}{(1 - \tau)r} \right\}. \]  

(41)

Using the same logic as in Li et al. (1997) and in proving Proposition 2 for the model without long-term debt, it can be shown that the firm should borrow the smallest amount that satisfies the liquidity constraint (8).

Corollary 2. For all \( Q \geq 0 \), it is optimal to let \( s_n b_n = 0, n = 1, \cdots, T \).

Two cases are consistent with this complimentary slackness condition and the constraints (8). One is \( s_n \geq 0 \) and \( b_n = 0 \) which is feasible but uninteresting because it reduces to a standard dynamic newsvendor problem. The other is \( s_n = 0 \) and \( b_n \geq 0 \) which we assume henceforth.

We define
\[ m_n = \frac{b_n}{(1 - \tau) + Q}. \]

(42)

Thus, \( b_n = (1 - \tau)(rm_n - Q) \), so (13) is the special case of (42) with \( Q = 0 \). With \( s = 0 \), it is apparent from (42) that \( m \) and \( b \) determine each other. It is convenient now to let \( m \) and \( y \) be the decision variables and treat \( b \) as a dependent variable given by \( b = (1 - \tau)(rm - Q) \). Let \( q(m) = 1 - F(m) \).

In this notation, from (41),
\[ T = \inf \{ n : D_n \leq m_n \}, \quad F(m_n) = P \{ T = n \mid T \geq n - 1 \}, \quad \text{and} \quad q(m) = P \{ T > n \mid T \geq n - 1, m_n = m \}. \]

Upon bankruptcy, the sequence of events is the same as in \$2.2 \) except that the tax collector is senior to the holder of the long-term bond who is senior to all other claimants. Hence, using the same notation as in \$2.2 \), the bondholder receives \( \min \{ \mu(Q), \mathcal{W}^+ \} \) and the “bank” that issued the short-term loan receives
\[ \mathcal{I} = \min \{ b_T, [\mathcal{W} - \mu(Q)]^+ \}. \]

The shareholders receive the residual which is
\[ \mathcal{J} = (\mathcal{W} - \mathcal{I})^+ = (\mathcal{W} - b_T - \mu(Q))^+ = ((1 - \tau)[rD_T + c_S(y_T - D_T)^+ - \mathcal{L}] + s_T - b_T - \mu(Q))^+. \]

The same steps that lead from (15) to (20) and which justify (35)-(37) can be used here with (39) and (40):
\[ \mathcal{V} = \Sigma_{n=1}^T \beta^{n-1}v_n + \beta^{T-1}\mathcal{J} \]
\[ = w_1 + (1 - \tau)c x_1 \]
\[ + \Sigma_{n=1}^{T-1} \beta^{n-1}\{(b_n - s_n)(1 - \beta) + (1 - \tau)[(\beta r - c)y_n - \beta(r + h - c)(y_n - D_n)^+ - \rho_n - Q]\} \]
\[ + \beta^{T-1}[b_T - s_T - (1 - \tau)(\rho_T + cy_T + Q)] \]
\[ + \beta^{T-1} \left[ (1 - \tau) [c_s(y_T - D_T)^+ + rD_T - \mathcal{L}] + s_T - b_T - \mu(Q) \right]^+. \quad (43) \]

By analogy with (20b), (21), and (22), we assume that there are sufficiently large fixed bankruptcy costs so that \( \{c_s(y_T - D_T)^+ + rD_T - \mathcal{L}\}^+ = 0 \) in (43). Therefore, as in (23),

\[ \frac{\mathcal{V}}{1 - \tau} = \frac{w_1}{1 - \tau} + cz_1 + \sum_{n=1}^{T-1} \alpha_1[y_n, b_n, s_n, D_n] - Q] + \beta^{T-1} \alpha_2[y_T, b_T, s_T, D_T] - Q], \quad (44) \]

where \( \alpha_1 \) and \( \alpha_2 \) are defined in (20a) and (20b).

Since \( w_1 = \eta + \mu(Q) \) and \( s_n = 0 \) for all \( n \),

\[ \mathcal{E}(\mathcal{V}) = \frac{\eta + \mu(Q)}{1 - \tau} + cz_1 + E \left( \sum_{n=1}^{\infty} \beta^{n-1} \left[ \Pi_{k=1}^{n-1} q(m_k) \right] \mathcal{K}(y_n, m_n; Q) \right). \quad (45) \]

So corresponding to (28), (29a), and (29b),

\[ \mathcal{K}(y, m; Q) = q(m) \int_{\infty}^{m} \alpha_1[y, (1 - \tau)(rm - Q), 0, u] - Q] dF(u) \]
\[ + F(m) \int_{0}^{m} \alpha_2[y, (1 - \tau)(rm - Q), 0, u] - Q] dF(u) \]
\[ = q(m) \int_{m}^{\infty} \alpha_1[y, (1 - \tau)(rm - Q), 0, u] dF(u) \]
\[ + F(m) \int_{0}^{m} \alpha_2[y, (1 - \tau)(rm - Q), 0, u] dF(u) \]
\[ - \{q(m)^2 + [F(m)]^2\}Q. \quad (46) \]

In \( \S 7 \) we give a detailed rationale for the constraint \( 0 \leq m \leq m^* \) in \( \S 4 \). Here, we impose a slightly stricter constraint, \( 0 \leq m \leq m^{**} \), where

\[ m^{**} = \min \left\{ F^{-1} \left( \frac{\beta}{2 + \beta} \right), m^* \right\} = \min \left\{ F^{-1} \left( \frac{\beta}{2 + \beta} \right), F^{-1} \left( \frac{\beta h + (1 - \beta)c}{2\beta h + (2 - \beta)c} \right) \right\}. \]

In \( \S 7 \) we show that \( m^* = m^{**} \) unless the discount factor \( \beta \) is absurdly low.

The following dynamic program corresponds to maximizing \( \mathcal{E}(\mathcal{B}) \):

\[ \phi(x; Q) = \sup \{ J(y, m; Q) : x \leq y, 0 \leq m \leq m^{**} \}, \quad (47a) \]
\[ J(y, m; Q) = \mathcal{K}(y, m; Q) + \beta q(m) \int_{m}^{\infty} \phi([y - u]^+; Q] dF(u). \quad (47b) \]

Extend the notation in (30) and define

\[ L(y, m, Q) = \frac{\mathcal{K}(y, m, Q)}{1 - \beta + \beta F(m)}. \quad (48) \]

and for each \( Q \geq 0 \) let \( [Y^C(Q), M^C(Q)] \) globally maximize \( L(\cdot, \cdot, Q) \) subject to the constraint \( 0 \leq m \leq m^{**} \). Similarly, let \( M^U(Q) \) globally maximize \( L(y^U, \cdot, Q) \) subject to the constraint \( 0 \leq m \leq m^{**} \).

Let \( B^C(Q) = (1 - \tau)rM^C(Q) \) and \( B^U(Q) = (1 - \tau)rM^U(Q) \).

The model now is essentially the same as that in \( \S 2 \) except for the coupon payment at the end of each period and the insignificantly tighter constraint with \( m^{**} \) instead of \( m^* \). So the argument in \( \S 3.4 \) leads to the following analogue of Proposition 1:
Corollary 3. \((y_n, m_n) = (Y^C(Q), M^C(Q)), n = 1, \cdots, T,\) is an optimal policy with respect to the set of initial states \(S = (-\infty, Y^C(Q)],\) and the consequent lifetime of the firm, \(T,\) has a geometric distribution with parameter \(q[M^C(Q)].\)

So if the initial inventory level, \(x_1,\) is not excessive (no greater than \(Y^C(Q)),\) then there is a myopic policy that is optimal. That is, \((y_n, m_n) = (Y^C(Q), M^C(Q))\) for all \(n\) is optimal.

Corollary 4. For each \(Q \geq 0,\) \(y^U \leq Y^C(Q)\) and \(B^U(Q) \leq B^C(Q).\)

However much long-term debt is included in the capital structure, the firm that coordinates its short-term decisions should have a higher target inventory level and borrow more in the short-term than its decentralized counterpart.

In summary, the properties of an optimal policy in §4 remain valid whether or not there is long-term debt in the capital structure.

In the following subsections, first we examine the dependence on the amount of long-term debt of the optimal dividend, short-term borrowing, and inventory decisions. Second, we seek an optimal amount of long-term debt.

6.2. Effects of the Amount of Long-Term Debt

In this section we establish that the amount of long-term debt (equivalently, the size of the coupon payment, \(Q\)) and the short-term decisions that maximize the market value of the firm are interdependent. Thus, in the notation used in §2.3 and earlier in this section, the maximal value of \(E(V),\) depends on \(Q.\) The value of \(Q\) that achieves the maximal EPV is examined in the next subsection.

Proposition 4. (i) Under assumption (50) (cf. §7), \(Y^C(\cdot), M^C(\cdot),\) and \(M^U(\cdot)\) are non-decreasing, and

(ii) \([B^C(Q + \Delta) - B^C(Q)]/\Delta \geq -(1 - \tau)\) and \([B^U(Q + \Delta) - B^U(Q)]/\Delta \geq -(1 - \tau)\) \((\Delta > 0).\) So the one-sided derivatives satisfy \(B^C(\cdot) \geq -(1 - \tau),\) \(B^U(\cdot) \geq -(1 - \tau),\) \(M^C(\cdot) \geq 0,\) and \(M^U(\cdot) \geq 0.\)

The proposition asserts that the inventory level and long-term debt are economic complements in the coordinated regime. That is, if \(Q\) were to be replaced by \(Q + \Delta\) with \(\Delta > 0,\) then it would be optimal to replace \(y = Y^C(Q)\) with the (weakly) greater \(y = Y^C(Q + \Delta) \geq Y^C(Q).\) Although \(M^C(Q)\) and \(Q\) are economic complements, it does not follow that \(B^C(Q)\) and \(Q\) have that property too. Since \(B^C(Q) = (1 - \tau)M^C(\cdot) - Q,\) we can conclude only that \(B^C(Q + \Delta) + (1 - \tau)\Delta \geq B^C(Q).\) Similarly, in the uncoordinated regime, \(B^U(Q + \Delta) + (1 - \tau)\Delta \geq B^U(Q).\)
6.3. Optimal Level of Long-term Debt

In this subsection we identify conditions that imply that a firm that coordinates its short-run decisions should be more heavily leveraged than one which decentralizes those decisions.

Let $\mathcal{V}_C(Q)$ denote the present value of the dividends when the initial capital structure entails a periodic coupon payment $Q$, and when decisions are coordinated via $(y_n, m_n) = [y^C(Q), m^C(Q)]$ for all $n$. Let $V_C(Q) = E[\mathcal{V}_C(Q)]$. Similarly, let $\mathcal{V}_U(Q)$ denote the present value of the dividends when decisions are uncoordinated, and let $V_U(Q) = E[\mathcal{V}_U(Q)]$. Let $Q^C$ and $Q^U$ maximize $V_C(\cdot)$ and $V_U(\cdot)$, respectively. For a fixed capital structure, i.e., for a fixed coupon payment $Q$, $V_C(Q)$ is the maximal value of an optimization problem in which $(y, m) = [y^U(Q), m^U(Q)]$ is feasible. Therefore, $V_U(Q) \leq V_C(Q)$ for each $Q$ and $V_U(Q^U) \leq V_C(Q^C)$. That is, the coordinated decisions yield a higher firm value than the decentralized ones. The principal result is that, if the short-term interest rate does not depend on the amount borrowed, i.e., if $\rho(\cdot)$ is affine, then the coordinated firm should be more highly leveraged.

**Proposition 5.** If $\rho(\cdot)$ is an affine function, then $Q^U \leq Q^C$.

A firm that coordinates its short-term decisions should incur greater long-term debt but, as we now show, not necessarily greater short-term debt than its decentralized counterpart. From Propositions 4 and 5 and definition (42) of $m_n$,

$$B^U(Q^U) \leq B^C(Q^U) = (1 - \tau)[rM^C(Q^U) - Q^U] \leq (1 - \tau)[rM^C(Q^C) - Q^U]$$

$$= [rM^C(Q^U) - Q^U] + (1 - \tau)(Q^C - Q^U) = B^C(Q^C) + (1 - \tau)(Q^C - Q^U).$$

So instead of $B^U(Q^U) \leq B^C(Q^C)$ we have $B^U(Q^U) \leq B^C(Q^C) + (1 - \tau)(Q^C - Q^U)$.

7. Justifications for Key Assumptions

In this section we provide rationales for some of the key financial and operational assumptions in earlier sections.

7.1. Key Financial Assumptions

The dominant assumption throughout the economics and finance literatures is that lending money is a passive and anonymous activity. While borrowers maximize their expected profit, in a competitive credit market, the best that lenders can hope to achieve is to be compensated fairly for credit risk without transferring a positive surplus from the borrowers. In this section, we establish that these assumptions imply (1) that there is an upper bound on short-term borrowing, and (2) that the total cost of borrowing is convex and increasing in loan size.
The Cost of Borrowing. The lenders know that in the event of bankruptcy they may lose some or all of the loan principle. So the amount that they charge the firm, $\rho(b)$, includes a default premium. Under standard assumptions in financial economics (cf. Tirole (2006), p. 116), the expected value of the net transfer from the lenders to the borrower is zero.

Assumption 1. Lenders' Condition. For each potential loan $b \geq 0$, the interest $\rho(b)$ charged to the borrower assures that the lender breaks even on average. It follows that:

$$q(m) \int_{m}^{\infty} [(1-\tau)\rho(b) + b] dF(u) + F(m) \int_{0}^{m} \mathcal{W}dF(u) = b,$$

where $q(m) = 1 - F(m)$. \hfill (49)

The first term on the left side of (49) is the expected revenue when the firm survives, namely, the post-tax loan principle and interest, and the second term is the expected residual claim available to the lender when the firm defaults and is liquidated (where $\mathcal{W}$ is the proceeds available for distribution in the event of bankruptcy given by (14) with $s_T = 0$). So borrowing costs increase with loan size. With IFR (increasing failure rate) demand, the borrowing cost function is convex and increasing.

Lemma 3. If the demand distribution function, $F(\cdot)$, is IFR, then $\rho(\cdot)$ is convex and increasing.

An analogous argument (cf. Leland (1994)) applied to the model in §6.1, leads to the conclusion that the amount of long-term debt that is associated with a periodic coupon payment of $Q$, $\mu(Q)$, is a concave and increasing function.

A Bound on the Marginal Cost of Borrowing. Although the "firm," namely the equity holders here, is not directly concerned with costs imposed on its lenders, it should be attentive to the impact of those costs on its own cost of borrowing, $\rho(b)$. When a firm maximizes the value of equity, as is done here, it is insufficient to ask whether a project generates value that exceeds its cost – a criterion adopted in the traditional NPV analysis. Instead, the firm's decision makers should ask whether the wealth the firm creates accrues to the equity holders or to the firm's lenders. Clearly, the adoption of projects that transfer wealth from the equity holders to the bank adversely affects the value of the firm's equity. The following condition is consistent with this perspective:

Assumption 2. Borrower's Condition. The firm will not make a short-term loan if its marginal cost of borrowing exceeds the marginal income generated by an investment that earns a required rate of return; equivalently,

$$1 - \beta \geq \rho'(b) \quad \text{for all } b > 0.$$ \hfill (50)
We refer to this assumption as the *borrower's condition* because it protects the borrower from "excessive" borrowing costs. Assumptions 1 and 2 are standard in the financial economics literature; Assumption 2 would be unnecessary if we were optimizing total firm value (value of equity plus value of debt) instead of shareholder value (cf. Grinblatt and Titman (2002)).

**An Imputed Bound on Short-Term Borrowing.** The lenders' and borrower's conditions (Assumptions 1 and 2) and \((1 - \tau)r \geq 1\) impute a bound on short-term borrowing, \(b\), and since \(b = (1 - \tau)rm\) from (27), also on \(m\). The latter yields a bound on the default probability that we state for the model without long-term debt. There is an analogous bound with long-term debt.

**Lemma 4.** At an optimal solution,

\[
F(m) \leq \frac{1 - \beta}{2 - \beta}. \tag{51}
\]

**The Bound** \(m \leq m^*\). The constraint \(m \leq m^*\), equivalently

\[
b = (1 - \tau)rm \leq (1 - \tau)rm^*, \tag{52}
\]

where

\[
m^* = F^{-1}\left[ \frac{\beta h + (1 - \beta)c}{2[\beta h + (2 - \beta)c]} \right], \tag{53}
\]

is conveniently imposed in §4, and here we show that it is redundant.

A firm enters bankruptcy in period \(n\) if it has survived periods \(1, 2, ..., n - 1\) and if \(D_n \leq m = b/[(1 - \tau)r]\). So the (conditional) default probability is \(F(m)\) which is bounded above by

\[
F(m^*) = \frac{\beta h + (1 - \beta)c}{2[\beta h + (2 - \beta)c]}.
\]

The condition (52) is convenient in proving Proposition 3, but it is redundant for all discounts factors \(\beta\) "close to 1", namely if

\[
\frac{3c}{2(c - h)} - \frac{1}{2} \sqrt{ \frac{c^2 + 8ch}{(c - h)^2} } \leq \beta \leq 1.
\]

The left inequality implies

\[
F(m) \leq \frac{1 - \beta}{2 - \beta} \leq \frac{\beta h + (1 - \beta)c}{2[\beta h + (2 - \beta)c]} = F(m^*),
\]

so (52) is necessarily satisfied due to Lemma 4.

**The Bound** \(m \leq m^{**}\). In §6 we constrain \(0 \leq m \leq m^{**}\) where

\[
m^{**} = \min \left\{ F^{-1}\left( \frac{\beta}{2 + \beta} \right), m^* \right\} = \min \left\{ F^{-1}\left( \frac{\beta}{2 + \beta} \right), F^{-1}\left( \frac{\beta h + (1 - \beta)c}{2[\beta h + (2 - \beta)c]} \right) \right\}.
\]
Now we show that $m^* = m^{**}$ unless the discount factor $\beta$ is absurdly low.

Notice that $m^* = m^{**}$ if

$$\frac{\beta h + (1 - \beta)c}{2\beta h + (2 - \beta)c} \leq \frac{\beta}{2 + \beta},$$

which corresponds to

$$0 \leq 2\beta c - (2 - \beta)(\beta h + c).$$

So $m^* = m^{**}$ if

$$0 \leq [2\beta - (2 - \beta)(\beta \gamma + 1)]c \text{ or equivalently } 0 \leq 2\beta - (2 - \beta)(\beta \gamma + 1).$$

The right inequality is necessarily satisfied if $\beta \geq 2/3$. If the calendar length of a “period” is a quarter of a year or smaller, then $\beta$ will be close to unity. For example, if $\beta = 0.9$ and $\gamma = 0.1$, then

$$2\beta - (2 - \beta)(\beta \gamma + 1) = 0.6.$$ In summary, $m^* = m^{**}$ unless the discount factor $\beta$ is absurdly low.

Observe that the probability of default is less than one-third because $m \leq m^{**} \leq F^{-1}[\beta/(2 + \beta)]$ and $\beta \leq 1$ imply $F(m) \leq \beta/(2 + \beta) \leq 1/3$.

7.2. Key Operational Assumptions

i.i.d. Demands. Except in §5, we assume that the demands $D_1, D_2, ...$ comprise a sequence of independent and identically distributed non-negative random variables. Integer-valued demands would have that property if the demands in successive periods were the result of discretizing a Poisson process. The Palm-Khintchine Theorem states that the aggregation of a large number of independent small-intensity equilibrium renewal processes is asymptotically a Poisson process. The behavioral interpretation is that the demand assumption is consistent with having a large number of customers having independent equilibrium renewal demand processes with low intensities. The classical reference is Khintchine (1969) which completed C. Palm’s earlier partial proof and first appeared in Russian in 1955. See §5.8 in Heyman and Sobel (2004b) for an exposition, and Lariviere and Van Mieghem (2004) for a recent re-discovery.

Excess Demand is Lost. The inventory dynamical equation is (3) where $x_{n+1} = (y_n - D_n)^+$. So we assume that demand which exceeds the supply of goods is lost, an assumption which is most reasonable in a model of a firm with retail customers. Instead, we could have assumed that excess demand is backordered, a reasonable assumption in a model of a firm with wholesale customers. The results would not have changed, but the model would have been slightly more complicated. See Hu et al. (2010) for modeling details.
8. Conclusions

We formulate and investigate a dynamic stochastic model of a firm's short-term operational and financial decisions. These decisions have interaction effects because both of them affect the working capital. The criterion in the model is to maximize the market value of the equity in the firm, namely the expected present value of the dividends. Under an assumption of limited liability at bankruptcy, i.e., the equity holders would lose control if bankruptcy were to occur, we compare the optimal short-run decisions when the firm coordinates financial and operational decisions with the decisions that should be made if the firm decentralized operational and financial decisions. Each period the operational decision is the amount to produce and add to inventory, and the financial decisions are the amounts to borrow short-term and the amount of a dividend to issue.

The model has a myopic policy that is optimal, and that permits us to show, regardless of whether the capital structure is all-equity or includes long-term debt, that the inventory levels, short-term loans, and market value of the equity are higher under coordination than decentralization. That is, coordination induces more aggressive operational and financial decisions and, consequently, greater exposure to risk than would be optimal if the firm were decentralized. That increases the market value of the equity. These conclusions could not be reached if the model were static instead of dynamic.

The optimality of a myopic policy facilitates the analysis, but previous studies of coordination had come to fruition only with an assumption of unlimited liability at bankruptcy (or "reorganization" bankruptcy). At first, the results are obtained under the assumption that product demands are independent and identically distributed random variables. Then we consider a non-stationary demand process and, surprisingly, find that no myopic policy could be optimal. Nevertheless, we reach conclusions regarding qualitative properties of optimal policies.

The degree of leverage, i.e., the amount of long-term debt, has interesting consequences. The combination of limited liability and tax benefits induces the coordinating firm with higher long-term debt to set a higher goods stock level and to secure larger short-term loans than its coordinating counterpart with lower long-term debt. So a more highly leveraged firm should take greater risks with its short-term decisions.

Since the capital structure affects short-term decisions including dividends, it also influences the market value of the equity. Thus, we ask: what degree of leverage maximizes the market value of the equity? We find that the coordinated firm should be more highly leveraged (in terms of long-term debt) than if it were decentralized.
References


Referee Appendix

Appendix A: Non-Stationary Demand

A.1. Demand Model

We employ the following demand model. Let

\[ D_n = \sum_{k=1}^{K} \theta_k D_{n-k} + \epsilon_n, \]  \hfill (54)

where \( \{\theta_k\} \) are known scalars, and \( \{\epsilon_n\} \) are independent, identically distributed, and non-negative random variables. Let \( \epsilon \) be a random variable with the same distribution as \( \epsilon_1 \), and let \( F(\cdot) \) denote the distribution function of \( \epsilon \).

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 \( \nu_n = \theta D_{n-1} \) reduces (54) to the following demand model used in the remainder of the paper:

\[ D_n = \theta D_{n-1} + \epsilon_n = \nu_n + \epsilon_n, \]  \hfill (55)

We interpret \( \nu_n \) as the deterministic part of demand in period \( n \).

A.2. Coordinated Operational and Financial Decisions

With demand process given by (55) and, as in §3, letting \( b = (1 - \tau)rm \), the dynamic program (25) can be written as

\[ \phi(x, \nu) = \sup \{ J(y, m, \nu) : m - \nu \geq 0, y \geq x \}, \]

\[ J(y, m, \nu) = \mathcal{K}(y, m, \nu) + \beta q(m - \nu) \int_{m - \nu}^{\infty} \phi((y - \nu - u)^+, \theta \nu + \theta u) dF(u), \]

where \( \mathcal{K} \) is given by (28) with \( u = \nu + \epsilon \). For reasons that become clear in the next paragraph, it is now convenient to rename \( x, y, \) and \( m \) as \( x', y', \) and \( m' \). The above dynamic program then becomes:

\[ \phi(x', \nu) = \sup \{ J(y', m', \nu) : m' - \nu \geq 0, y' \geq x' \}, \]  \hfill (56a)

\[ J(y', m', \nu) = \mathcal{K}(y', m', \nu) + \beta q(m' - \nu) \int_{m' - \nu}^{\infty} \phi((y' - \nu - u)^+, \theta \nu + \theta u) dF(u), \]  \hfill (56b)

which can be simplified by letting (see Lamond and Sobel (1995) for a similar transformation)

\[ x = x' - \nu, \]  \hfill (57a)

\[ y = y' - \nu, \]  \hfill (57b)

\[ m = m' - \nu. \]  \hfill (57c)

In words, the "non-prime" variables on the left side of (57) represent the amounts of inventory, \( x' \) and \( y' \), and a bankruptcy threshold, \( m' \), that are in excess of the deterministic part of demand \( \nu \).

Equation (56), the transformation (57), and a little algebra then yield a transformed dynamic program

\[ \phi(x, \nu) = (r - c)\nu + \sup \{ J(y, m, \nu) : m \geq 0, y \geq x \}, \]  \hfill (58a)
\[ J(y, m, \nu) = K(y, m) + \beta q(m) \int_{m}^{\infty} \phi((y - u)^+, \theta \nu + \theta u) dF(u), \]  

(58b)

where \( K(y, m) \) is given by (28), \( q(m) = 1 - F(m) \), and \( F \) is the distribution function of \( \varepsilon \), the stochastic portion of demand, \( D \), given by (55). \textit{Note that ultimately, we are interested in characterizing an optimal policy for the dynamic program (56). However, to do this, we study an optimal policy in the dynamic program (58) and then recover a policy optimal in (56) using (57).}

In §3, where demands \( D_1, D_2, \ldots \) are i.i.d. non-negative random variables, we establish that a policy that maximizes the expected present value of dividends, \( E(\mathcal{V}) \), is obtained by solving the dynamic program (25), and that this policy is myopic. We find this myopic policy by maximizing the function \( L(y, m) \) given by (30). That is, we find \((y, m) = (y^C, m^C)\) optimal in

\[ \max_{y, m} L(y, m) = \left\{ \frac{K(m, y)}{(1 - \beta q(m))} : m \geq 0, y \geq 0 \right\}. \]  

(59)

Here, demands \( D_1, D_2, \ldots \) follow AR(1) process and a policy that maximizes the expected present value of dividends, \( E(\mathcal{V}) \), is obtained by solving the dynamic program (56). This is done first by solving the dynamic program (58) and then (using (57)) letting

\[ y' = y + \nu, \]

\[ m' = m + \nu. \]

If the dynamic program (58) had a myopic policy, then this policy would be established by finding \((y, m) = (y^{U}, m^{U})\) optimal in

\[ \max_{y, m} L_{\nu}(y, m, \nu) = \left\{ \frac{K(m, y) + (r - c)\nu}{(1 - \beta q(m))} : m \geq 0, y \geq 0 \right\}. \]  

(60)

However, as it is readily seen from (60), \( (y^{U}, m^{U}) \) depend on \( \nu \), and it is easy to construct an example where the action \((y_n, m_n) = (y^{U}, m^{U})\) is not feasible in all periods \( n \). The next set of results, however, constructs lower and upper bounds on \((y^{U}, m^{U})\), which do \textit{not} depend on \( \nu \). This allows us to conclude that a policy optimal in the dynamic program (56) has monotone lower and upper bounds that are linear in \( \nu \).

**Lemma 5.** Let \((y, m) = (y^C, m^C)\) and \((y, m) = (y^{U}, m^{U})\) be optimal in (59) and (60) respectively. Then

(i) \((y^{U}, m^{U}) \geq (y^C, m^C)\);

(ii) \((y^{U}, m^{U}) \leq (y(m^*), m^*)\), where \( m^* \) is given by (34);

(iii) In (58), the action \((y_n, m_n) = (y^C, m^C)\) is feasible for all \( n \).

**Proof of Lemma 5.** First, showing part (i). Using (59), \((y, m) = (y^C, m^C)\) solves

\[ L^{(1)}(y, m) = 0 \quad \text{and} \quad L^{(2)}(y, m) = 0. \]  

(61)

Similarly, using (60), \((y, m) = (y^{U}, m^{U})\) solves

\[ L^{(1)(y, m)} = L^{(1)}(y, m) = 0 \quad \text{and} \quad L^{(2)}(y, m) = L^{(2)}(y, m) + \beta(r - c)\nu q^{(2)} = 0. \]  

(62)
Now substituting \((y, m) = (y^C, m^C)\) into (62) reveals
\[
L_u^{(1)}(y, m) = L_u^{(1)}(y, m) \geq 0 \quad \text{and} \quad L_u^{(2)}(y, m) = L_u^{(2)}(y, m) + \beta(r - c)\nu \geq 0. \tag{63}
\]
Concavity (Lemma 6) and supermodularity (Lemma 2) of \(L_u\) imply \((y^C, m^C) \geq (y^C, m^C)\). To establish part (ii), note that from part (i) both \(y^C\) and \(m^C\) increase in \(\nu\); the assumption \(m^C \leq m^*\) implies the result.
Next, showing part (iii). If \(y_n = y^C\) for some \(n\), then \(x_{n+1} = (y_n - \epsilon)^+ = (y^C - \epsilon)^+ \leq y^C\) because \(\epsilon \geq 0\). Hence \(y_{n+1} = y^C\) is feasible. Since all short-term debt, \(b\), is fully repaid at the end of each period \(n\), and \(b = r(1 - \tau)m\), then \(m_{n+1} = m^C\) is also feasible for all \(n\). \(\square\)
Lemma 5 and the transformation (57) now imply \((y, m)\) optimal in (56) have upper and lower bounds that are linear in \(\nu\). In particular,
\[
y^C \in \nu + [y^C, y(m^*)], \tag{64a}
m^C \in \nu + [m^C, m^*]. \tag{64b}
\]
The bounds on an optimal coordinated borrowing policy obtain from (27), which implies \(\nu^C = (1 - \tau)rm^C\). Together with (64b), this yields:
\[
\nu^C \in (1 - \tau)\nu + [(1 - \tau)rm^C, (1 - \tau)rm^*].
\]
Therefore an optimal coordinated policy induces bounds on physical goods level and short-term loan that increase linearly with the deterministic part of demand. The respective growth rates are 1 and \(r(1 - \tau)\).

A.3. Uncoordinated Operational and Financial Decisions

The analysis is very similar to the coordinated case presented in the previous section and therefore we skip some details. The same steps that led to the uncoordinated policy \((y^U, m^U)\) in §4.1, and the transformation (57) now reveal that an uncoordinated policy that maximizes the expected present value of dividends when demands \(D_1, D_2, \ldots\) follow AR(1) process takes the action
\[
y^U = \nu + y^U, \tag{65a}
m^U = \nu + m^U, \tag{65b}
\]
in every period \(n = 1, 2, \ldots\). The supply level \(y = y^U\) is optimal in
\[
\max_{\nu} \quad (\beta r - c)\nu - \beta(r + h - c)E((y - \epsilon)^+),
\]
and \(m = m^U\) is optimal in
\[
\max_{m} \quad L_u(y, m, \nu) = \left\{ \frac{\hat{K}(m, y) + (\tau - c)\nu}{(1 - \beta q(m))} : m \geq 0 \bigg| y = y^U \right\}.
\]
As in the coordinated case, an optimal uncoordinated borrowing policy obtains from (27), which implies \(\nu^U = (1 - \tau)rm^U\). Together with (65b), this yields:
\[
\nu^U = r(1 - \tau)(\nu + m^U).
\]
The expressions for \(y^U\) (given by (65a)) and \(b^U\) (above) therefore reveal that an optimal uncoordinated policy induces a physical goods base-stock level and short-term loan that increase linearly with the deterministic part of demand. Analogously to the coordinated case, the respective growth rates are 1 and \(r(1 - \tau)\).
A.4. Comparison of Coordinated with Uncoordinated Solutions

Finally, (64) and (65) yield that Proposition 3 established in §4 remains valid when demands $D_1, D_2, \ldots$ follow AR(1) process.

**Proposition 6.** Simultaneous optimization of borrowing level, $b$, and supply level, $y$, results in higher optimal level of borrowing and higher optimal level of inventory. That is: (i) $y^{C} \geq y^{U}$ and (ii) $b^{C} \geq b^{U}$.

**Proof of Proposition 6.** Part (i) is immediate from (64a) and (65a), since from Proposition 3 we have $y^{C} \geq y^{U}$. Adding $\nu$ to both sides of the inequality yields $y^{C} \geq y^{U}$. Part (ii) is established by showing that a policy in (58) that chooses $(y_n, m_n) = (y^C, \bar{m})$, where $m = \bar{m}$ solves (60) with $y = y^{C}$ is feasible in period $n$, and that $(y^{C^U}, m^{C^U}) \geq (y^{C}, \bar{m}) \geq (y^{U}, m^{U^*})$. The proof is analogous to that of Lemma 5 parts (i) and (iii). □

Appendix B: Lemmas and Proofs

**Proof of Proposition 2.**

Let $t = b - s$ and define

$$
\bar{q}(t) = P\{ (1 - \tau)D > t \},
\tag{66}
$$

$$
\bar{m}(t) = \frac{t}{(1 - \tau)r}.
\tag{67}
$$

The constraints in the dynamic program are $b \geq 0, s \geq 0$, and $y \geq x$ which corresponds to $s + t \geq 0, s \geq 0$, and $y \geq x$, which can be written as $s \geq (-t)^+$ and $y \geq x$. So the dynamic program can be written as

$$
\phi(x) = \sup\{ J(y, s + t, s) : s \geq (-t)^+, y \geq x \},
\tag{68a}
J(y, s + t, s) = K(y, s + t, s) + \beta \bar{q}(t) \int_{\bar{m}(t)}^{\infty} \phi((y - u)^+) dF(u).
\tag{68b}
$$

The joint optimization with respect to $y, t,$ and $s$ can be nested, so

$$
\phi(x) = \sup_{n,t,y} \{ J(y, s + t, s) : s \geq (-t)^+, y \geq x \},
= \sup_{y, t} \{ \sup_{s} \{ J(y, s + t, s) : s \geq (-t)^+ \} : y \geq x \},
= \sup_{y, t} \{ \sup_{s} \{ K(y, s + t, s) : s \geq (-t)^+ \} + \beta \bar{q}(t) \int_{\bar{m}(t)}^{\infty} \phi((y - u)^+) dF(u) : y \geq x \}.
\tag{69}
$$

The remainder of the proof shows that $s = (-t)^+$ is optimal in

$$
\sup_{s} \{ K(y, s + t, s) : s \geq (-t)^+ \},
\tag{70}
$$

Therefore, $s = (-t)^+$ is optimal in (69). So in (25), it is optimal to let $s = (s - b)^+$, which corresponds to $bs = 0$. Definitions (20a), (20b), (24), (66), and (67) yield

$$
K(y, s + t, s) = \bar{q}(t) \int_{\bar{m}(t)}^{\infty} \alpha_1(y, s + t, s, u) dF(u) + [1 - \bar{q}(t)] \int_{0}^{\bar{m}(t)} \alpha_2(y, s + t, s, u) dF(u).
$$

Therefore,

$$
\frac{\partial K(y, s + t, s)}{\partial s} = \bar{q}(t) \int_{\bar{m}(t)}^{\infty} \left[ \alpha_2^{(2)}(y, s + t, s, u) + \alpha_1^{(3)}(y, s + t, s, u) \right] dF(u)
+ [1 - \bar{q}(t)] \int_{0}^{\bar{m}(t)} \left[ \alpha_2^{(2)}(y, s + t, s, u) + \alpha_2^{(3)}(y, s + t, s, u) \right] dF(u).
$$

We use superscript to denote partial derivative. So $\alpha^{(2)}_i$ denotes $\partial \alpha_i / \partial b$ and $\alpha^{(3)}_i$ denotes $\partial \alpha_i / \partial s$ ($i = 1, 2$). It follows from (20a) and (20b) that $\alpha^{(2)}_1 + \alpha^{(3)}_1 = \alpha^{(2)}_2 + \alpha^{(3)}_2 = -\rho' \leq 0$ and $\partial K(y, s + t, s) / \partial s = -\rho' \leq 0$. So (70) is achieved by the smallest feasible $s$, namely $s = (-t)^+$.

Proof of Lemma 1.

The proof establishes that $L^{(11)}(y, m) \leq 0$. Since from (30) we have

$$L^{(11)}(y, m) = \frac{K^{(11)}(y, m)}{1 - \beta F(m)},$$

and $1 - \beta F(m) > 0$ for all $m \geq 0$, then $L^{(11)}(y, m) \leq 0$ if and only if $K^{(11)}(y, m) \leq 0$.

The expression for $K(y, m)$ in (30) is given by (28). To obtain the expression for $K^{(1)}$ using (28), it is convenient to write $K(y, m)$ as follows:

$$K(y, m) = F(m) \int_{y}^{\infty} \alpha(y, (1 - \tau)r, m, 0, u | u > y) dF(u) + F(m) \int_{m}^{y} \alpha(y, (1 - \tau)r, m, 0, u | u \leq y) dF(u) + f(m) \int_{m}^{y} \alpha(y, (1 - \tau)r, m, 0, u) dF(u),$$

(71)

where we use the fact that $y > m$ (cf. §2.2).

Using (71), (29a), and (29b), the expression for $K^{(1)}$ is given as follows:

$$K^{(1)}(y, m) = F(m) \int_{y}^{\infty} \alpha^{(1)}(y, (1 - \tau)r, m, 0, u | u > y) dF(u) + F(m) \int_{m}^{y} \alpha^{(1)}(y, (1 - \tau)r, m, 0, u | u \leq y) dF(u) + f(m) \int_{y}^{\infty} \alpha^{(1)}(y, 1 - \tau)r, m, 0, u | u \leq y) dF(u)$$

$$+ f(m) \int_{y}^{\infty} \alpha^{(1)}(y, 1 - \tau)r, m, 0, u | u > y) dF(u)$$

$$= [1 - F(m)] F(y) - (\beta r - c) - \beta(r + h - c) - F^2(m) c$$

$$= F(m) (\beta r - c) - F(m) F(y) \beta(r + h - c) + F(m) F(m) [\beta h + c(1 - \beta)] - cF^2(m).$$

(72)

From (72),

$$K^{(11)}(y, m) = -\beta F(m) f(y) (r + h - c) \leq 0.$$

Proof of Lemma 2.

We show $L^{(12)}(y, m) \geq 0$ for $y \geq y^U$ and $0 \leq m \leq m^*$. From the definition of $L(\cdot, \cdot)$ in (30) and the specification of $K^{(1)}$ in (72),

$$L^{(12)}(y, m) = \frac{K^{(12)}(y, m)[1 - \beta + \beta F(m)] - \beta f(m) K^{(1)}(y, m)}{[1 - \beta + \beta F(m)]^2},$$

(73)

where

$$K^{(12)}(y, m) = [\beta h + c(1 - \beta)] [\beta f(m) F(m) + F(m) f(m)] + f(m) F(y) \beta(r + h - c)$$

$$- f(m) (\beta r - c) - 2cF(m) f(m)$$

$$= f(m) \left\{ [\beta h + c(1 - \beta)] - \beta(r + h - c) F(y) - (\beta r - c) - 2cF(m) \right\}. $$

(74)

If $L^{(12)}(y^U, m) \geq 0$ (0 \leq m \leq m^*), then $L^{(12)}(y, m) \geq 0$ for all $y \geq y^U$ because $y^U \leq y^C$ (part (i)) and $L^{(12)}(\cdot, m)$ is non-decreasing (for each $0 \leq m \leq m^*$). The latter property follows from $K^{(12)}(\cdot, m)$ non-decreasing (see the right side of (74)) and $K^{(1)}(\cdot, m)$ non-increasing (Lemma 1).
From (73),
\[ L^{(12)}(y, m)[1 - \beta + \beta F(m)]^2 = K^{(12)}(y, m)[1 - \beta + \beta F(m)] - \beta f(m)K^{11}(y, m). \]

So \( L^{(12)}(y', m) \geq 0 \) if
\[ K^{(12)}(y', m)[1 - \beta + \beta F(m)] - \beta f(m)K^{11}(y', m) \geq 0. \]

From (72) and \( F(y') = (\beta r - c)/[\beta(r + h - c)] \),
\[ K^{(12)}(y', m) = f(m)\left\{ \beta h + c(1 - \beta)\right\}[1 - 2F(m)] - 2cF(m) \]
whose non-negativity corresponds to
\[ F(m) \leq \frac{\beta h + (1 - \beta)c}{2[\beta h + (2 - \beta)c]} \]
which corresponds to \( m \leq m^* \), i.e., \( b \leq (1 - \tau)rm^* = u \). ❄️

Proof of Lemma 3.

Assume that \( \rho(\cdot) \) is differentiable, and let \( F^* = F(m), \tilde{F}^* = \tilde{F}(m), \tilde{f}^* = f(m), \) and \( W^* = c_s(y - m) + b/(1 - \tau) - L \), so \( W^* \geq W \) if \( D_r < m \). Taking the first-order derivative with respect to \( b \) on both sides of (49) yields
\[ (\tilde{F}^*)^2[(1 - \tau)\rho' + 1] - \frac{2\tilde{F}^* f^*}{(1 - \tau)r}[(1 - \tau)\rho + b] + \frac{f^*}{(1 - \tau)r} \int_0^m WdF(u) - \frac{F^* f^* W^*}{(1 - \tau)r} = 1. \]

So
\[ (\tilde{F}^*)^2(1 - \tau)\rho' = 1 - (\tilde{F}^*)^2 + \frac{2\tilde{F}^* f^*}{(1 - \tau)r}[(1 - \tau)\rho + b] - \frac{f^*}{(1 - \tau)r} \int_0^m WdF(u) + \frac{F^* f^* W^*}{(1 - \tau)r} \]
because
\[ \frac{f^*}{(1 - \tau)r} \int_0^m WdF(u) \leq \frac{F^* f^* W^*}{(1 - \tau)r}. \]

Therefore,
\[ (\tilde{F}^*)^2(1 - \tau)\rho' \geq 1 - (\tilde{F}^*)^2 + \frac{2\tilde{F}^* f^*}{(1 - \tau)r}[(1 - \tau)\rho + b] \geq 0. \]

Hence, \( \rho'(\cdot) \geq 0 \). Moreover,
\[ (1 - \tau)\rho' \geq \left( \frac{1}{(\tilde{F}^*)^2} - 1 \right) + \frac{2\tilde{F}^*}{\tilde{F}^* : \rho + b/(1 - \tau)}. \]

(75)

Note that the right side of (75) is non-negative for any \( m \geq 0 \), and it is increasing in \( m \) if \( f/\tilde{F} \) is increasing for the following argument: (1) the first term in the parenthesis on the right side is non-negative and increasing in \( m \), and (2) the second term on the right is also increasing in \( m \) since \( f/\tilde{F} \) is increasing in \( m \). By definition, \( m = b/r(1 - \tau) \), a greater \( m \) means greater amount of loan. Therefore, (75) implies that the cost of borrowing is an increasing convex function of the loan size. ❄️

Proof of Lemma 4. The left side of (50) and right side of (75) (see the proof of Lemma 3 in Appendix B) imply
\[ (1 - \beta) \geq \rho'(b) \geq (1 - \tau)\rho'(b) = (1 - \tau)\rho'((1 - \tau)rm) \geq (1 - \tau)\rho'(m). \]
\[
\begin{align*}
&\geq \left( \frac{1}{(F(m))^2} - 1 \right) + \frac{2f(m)}{F(m)} \cdot \frac{\rho + b/(1-\tau)}{r} \geq \left( \frac{1}{(F(m))^2} - 1 \right), \\
\text{or} \quad &1 - \beta \geq \left( \frac{1}{(F(m))^2} - 1 \right).
\end{align*}
\tag{76}
\]

Next, note that \(0 \leq F(m) \leq 1\) for all \(m\) implies
\[
\left( \frac{1}{(F(m))^2} - 1 \right) \geq \left( \frac{1}{F(m)} - 1 \right). 
\tag{77}
\]

Combine the left side of (76) with the right side of (77) to obtain
\[
(1 - \beta) \geq \left( \frac{1}{F(m)} - 1 \right) \quad \text{or} \quad F(m) \leq \frac{1 - \beta}{2 - \beta},
\]
which completes the proof. \(\square\)

**Lemma 6.** Let \(L\) be given by (30) and let \((y, m) = (y^C, m^C)\) be optimal in (59). Then \(L^{(22)}(y^C, m^C) \leq 0.\)

**Proof of Lemma 6.** Differentiating \(L\) with respect to \(m\) twice yields
\[
L^{(22)} = \left[ \frac{\mathcal{K}}{1 - \beta q} \right]^{(22)} = \frac{[\mathcal{K}^{(22)}(1 - \beta q) + \mathcal{K} q^{(22)}](1 - \beta q)^2 + 2(1 - \beta q) q^{(22)}[\mathcal{K}^{(2)}(1 - \beta q) + \mathcal{K} q^{(2)}]}{(1 - \beta q)^3}.
\tag{78}
\]

Because the first-order condition with respect to \(m\) (namely \(L^{(2)} = 0\)) requires \((1 - \beta q)\mathcal{K}^{(2)} + \beta \mathcal{K} q^{(2)} = 0\), the right side of (78) simplifies to:
\[
L^{(22)} = \left[ \frac{\mathcal{K}}{1 - \beta q} \right]^{(22)} = \frac{\mathcal{K}^{(22)}(1 - \beta q) + \mathcal{K} q^{(22)}}{(1 - \beta q)^2}.
\]

Therefore, it remains to show that at \((y, m) = (y^C, m^C)\) we have
\[
\mathcal{K}^{(22)}(1 - \beta q) + \mathcal{K} q^{(22)} \leq 0. 
\tag{79}
\]

Using Leibnitz' rule on (28) yields
\[
\mathcal{K}^{(2)} = -f(A_1 - A_2) + A_3, 
\tag{80}
\]
where
\[
A_1 = \int_m^\infty \alpha_1 dF(u) - \int_0^m \alpha_2 dF(u), \quad A_2 = F\alpha_2^* - \bar{F}\alpha_1^*, \quad A_3 = \bar{F} \int_m^\infty \alpha_1^{(2)} dF(u) + F \int_0^m \alpha_2^{(2)} dF(u).
\]

The expressions \(\alpha_1^*\) and \(\alpha_2^*\) are used to denote \(\alpha_1\) and \(\alpha_2\), given by (29a), (29b) evaluated at \(u = m\). Differentiating (80) again with respect to \(m\) yields:
\[
\mathcal{K}^{(22)} = -f'(A_1 - A_2) - f(A_1^{(2)} - A_2^{(2)}) + A_3^{(2)},
\]
where
\[
A_1^{(2)} = -\alpha_1^* f + \alpha_2^* f + \int_m^\infty \alpha_1^{(2)} dF(u) - \int_0^m \alpha_2^{(2)} dF(u), \quad A_2^{(2)} = (\alpha_2^* + \alpha_1^*) \left( f + f' \left( F\alpha_2^* - (1-F)\alpha_1^* \right) \right) + f' \left( F\alpha_2^* - (1-F)\alpha_1^* \right), \quad A_3^{(2)} = -2r(1 - \beta - (1-\tau)\rho') f(1-F) - r^2 (1-\tau)^2 \rho'' (1-F)^2 \\
+ 2r (1 - (1-\tau)\rho') F - r^2 (1-\tau)^2 \rho'' (F)^2.
\]
The expressions for $a_1^{(2)}, a_2^{(2)}, a_1^{(2)},$ and $a_2^{(2)}$ are given by:

\[
\begin{align*}
   a_1^{(2)} &= a_1^{(2)}, \\
   a_1^{(2)} &= -r \beta + a_2^{(2)}, \\
   a_2^{(2)} &= a_2^{(2)}, \\
   a_2^{(2)} &= r (1 - (1 - \tau) \rho').
\end{align*}
\]

Next, the borrower’s condition (50) and algebra imply $f^* A_1 \geq 0$, $A_1^{(2)} \geq 0$, $f^* A_2 \leq 0$, $A_2^{(2)} \leq 0$, and $A_3 \leq 0$. Consequently, $K^{(22)} \leq 0$, $K(y, m) \geq 0$ and $q^{(22)} \leq 0$ imply (79), which completes the proof. □

**Proof of Corollary 1.**

The total differentiation of the first-order condition, $L^{(1)} = 0$ (where $L$ is given by (30)), with respect to $y$ yields

\[
\frac{\partial y}{\partial \nu} = -\frac{K^{(12)}}{K^{(11)}}.
\]

Concavity, $K^{(11)} \leq 0$ (cf. Proof of Lemma 1), and supermodularity, $K^{(12)} \geq 0$ (cf. Proof of Lemma 2), now imply the result. □

**Proof of Proposition 4.**

Part (i) is a consequence of supermodularity of $L(\cdot, \cdot)$ which is established in Lemma 7. For part (ii), since $b = (1 - \tau)(\nu m - Q)$,

\[
\frac{b((\nu + \Delta) - b(\nu))}{\Delta} = (1 - \tau)[\nu m(\nu + \Delta) - (\nu + \Delta)] - (1 - \tau)[\nu m(\nu) - \nu m(\Delta)] = (1 - \tau)\left[\frac{\nu m(\nu + \Delta) - \nu m(\Delta)}{\Delta} - 1\right] \geq -(1 - \tau)
\]

with the inequality implied by the monotonicity of $m(\cdot)$ from part (i). □

**Lemma 7.** $L(\cdot, \cdot)$ is a supermodular function on $[y, \infty) \times [0, \nu] \times [0, \infty)$.

**Proof of Lemma 7.** By Lemma 2, $L^{(12)} \geq 0$ on $[y, \infty) \times [0, \nu]$. Therefore it remains to show (i) $L^{(13)} \geq 0$, and (ii) $L^{(23)} \geq 0$ on $[y, \infty) \times [0, \nu] \times [0, \infty)$. Part (i) is immediate since $L^{(13)}(y, m, Q) = K^{(13)}(y, m, Q)/(1 - \beta q(m))$ and $K^{(13)}(y, m, Q) = 0$. To show part (ii), note that

\[
L^{(23)}(y, m, Q) = \frac{K^{(23)}(y, m, Q)[1 - \beta q(m)] - \beta f(m)K(y, m, Q)}{[1 - \beta q(m)]^2}.
\]

Differentiating again with respect to $Q$ yields:

\[
L^{(23)}(y, m, Q) = \frac{K^{(23)}(y, m, Q)[1 - \beta q(m)] - \beta f(m)K^{(3)}(y, m, Q)}{[1 - \beta q(m)]^2}.
\]

Now, to show $L^{(23)} \geq 0$, it sufficient to establish that $K^{(23)} \geq 0$ and $K^{(3)} \leq 0$.

From (20a), (20b), and (46),

\[
K(y, m; Q) = q(m) \int_{m}^{\infty} \left\{ \frac{1 - \beta}{1 - \tau} (\nu m - Q) + (\beta r + c) (y - \beta (r + h - c) (y - u)^+ \right.
\]

\[
- \rho(1 - \tau)(\nu m - Q) \right\} dF(u)
\]

\[
+ F(m) \int_{m}^{\infty} \left\{ \frac{\nu m - Q}{1 - \tau} - \rho(1 - \tau)(\nu m - Q) - cy \right\} dF(u)
\]

\[
- \{q(m)^{2} + [F(m)]^{2}\} Q
\]

\[
= \{q(m)^{2} [1 - (1 - \beta)(\nu m - Q) - \rho(1 - \tau)(\nu m - Q)] + (\beta r + c) y
\]

\[- \rho(1 - \tau)(\nu m - Q) - cy\}
\]

\[
- \{q(m)^{2} + [F(m)]^{2}\} Q.
\]
Differentiating the above equation with respect to $Q$ yields:

$$
{\mathcal{K}}^{(3)}(y, m, Q) = [q(m)]^2 (1-\beta + (1-\tau)\rho'(1-\tau)(rm - Q)) + \frac{2f(m)}{} \left[ q(m) [(1-\beta + (1-\tau)\rho'(1-\tau)(rm - Q)] + F(m) \right] 
$$

Now, assumption (50) implies that the first two terms of (81) are non-positive, and so

$$
{\mathcal{K}}^{(3)}(y, m, Q) \leq 0.
$$

Next, from (81),

$$
{\mathcal{K}}^{(23)}(y, m, Q) = \frac{r(1-\tau)\rho''(1-\tau)(rm - Q)}{[q(m)]^2 + [F(m)]^2} + 2f(m) \left[ q(m) [(1-\beta + (1-\tau)\rho'(1-\tau)(rm - Q)] + F(m) \right].
$$

Since the first line of (82) is non-negative, combining terms in the second and third lines yields

$$
{\mathcal{K}}^{(23)}(y, m, Q) \geq 2f(m) \left[ -(1-\tau)\rho'(1-\tau)(rm - Q) + (1+\beta)q(m) - F(m) \right]
$$

which is non-negative for all $m \leq m^*$. Since $m^* \leq m^{**}$, (cf. §7) then the proof is complete.

\[\square\]

**Proof of Proposition 5.** Let $\mathcal{K}(Q)$ and $\mathcal{K}^{(j)}$ denote $\mathcal{K}[y(Q), m(Q), Q]$ and $\mathcal{K}^{(j)}[y(Q), m(Q), Q]$, $j = 1, 2, 3$ respectively. Using (46)

$$
\mathcal{K}'(Q) = \mathcal{K}^{(1)} \frac{dy}{dQ} + \mathcal{K}^{(2)} \frac{dm}{dQ} + \mathcal{K}^{(3)}.
$$

From (45),

$$
\frac{V'(Q)}{1-\tau} = \frac{\mu'(Q)}{1-\tau} + \frac{\mathcal{K}'(Q)[1 - \beta q(m)] - \beta f(m) \mathcal{K} dm/dQ}{[1 - \beta q(m)]^2}.
$$

Therefore,

$$
\frac{V'(Q)}{1-\tau} = \frac{\mu'(Q)}{1-\tau} + \left[ \frac{\mathcal{K}^{(1)} dx/dQ + \mathcal{K}^{(2)} dm/dQ + \mathcal{K}^{(3)}[1 - \beta q(m)] - \beta f(m) \mathcal{K} dm/dQ}{[1 - \beta q(m)]^2} \right].
$$

Since

$$
L(y, m, Q) = \frac{\mathcal{K}(y, m, Q)}{1 - \beta q(m)},
$$

and if $0 < m = M^C(Q) < m^{**}$ then $L^{(1)}(y, m, Q) = L^{(2)}(y, m, Q) = 0$ at $(y, m) = (Y^C(Q), M^C(Q))$,

$$
L^{(1)}(y, m, Q) = \frac{\mathcal{K}^{(1)}(y, m, Q)}{1 - \beta q(m)} = 0,
$$

$$
L^{(3)}(y, m, Q) = \frac{\mathcal{K}^{(2)}(y, m, Q)[1 - \beta q(m)] - \beta f(m) \mathcal{K}(y, m, Q)}{[1 - \beta q(m)]^2} = 0.
$$

Therefore, conclude that at $(y, m) = (Y^C(Q), M^C(Q))$ with $0 < M^C(Q) < \infty$,

$$
\mathcal{K}^{(1)}(y^C, m^C, Q) = 0 \quad \mathcal{K}^{(2)}(y^C, m^C, Q)[1 - \beta q(m)] - \beta f(m^C) \mathcal{K}(y^C, m^C, Q) = 0.
$$
If $M^C(Q) = 0$ or $M^C(Q) = m^{**}$, since $m^C$ is a constant, $dm^C/dQ = 0$. This follows because Lemma 7 implies \( \{Q : M^C(Q) = 0\} \) has the form \([0, Q_1]\), \( \{Q : 0 < M^C(Q) < m^{**}\} \) has the form \((Q_1, Q_2)\), and \( \{Q : M^C(Q) = m^{**}\} \) has the form \([Q_2, \infty)\).

Similarly, at \((y, m) = (y^U, m^U)\), if \(0 < m^U < m^{**}\), then \(L^{(2)}(y^U, m, Q) = 0\) yields

\[
K^{(2)}(y^U, m^U, Q)[1 - \beta q(m^U)] - \beta f(m^U)K(y^U, m^U, Q) = 0,
\]

but \(dm^U/dQ \neq 0\). If \(m^U = 0\) or \(m^U = m^{**}\), then \(dm^U/dQ = 0\). Moreover, \(y^U\) is also a constant that is independent from \(Q\), so \(dy^U/dQ = 0\), Hence, \(85\) becomes

\[
\frac{V'(Q)}{1 - \tau} = \frac{\mu'(Q)}{1 - \tau} + \frac{K^{(3)}(y, m, Q)}{1 - \beta q(m)} \quad (y, m) = [y^C(Q), m^C(Q)] \text{ or } (y, m) = [y^U, m^U(Q)].
\]

From \(81\), since \(\rho(\cdot)\) is affine,

\[
K^{(3)}(y, m, Q) = -(1 - \tau)^2 \rho''([1 - \tau]r \mathfrak{m} - Q) \{F(m)^2 + [q(m)]^2\} = 0.
\]

So \(V_C(\cdot)\) and \(V_U(\cdot)\) are concave on their domain when \((y, m) = [y^C(Q), m^C(Q)]\) and \((y, m) = [y^U, m^U(Q)]\), respectively. Therefore,

\[
V_U(Q^U)|_{(y, m) = [y^U(Q^U), m^U(Q^U)]} = 0,
\]

which is equivalent to

\[
\mu'(Q^U) + (1 - \tau)\frac{K^{(3)}[y^U, m^U(Q^U)]}{1 - \beta q(m^U(Q^U))} = 0.
\]

Because of the concavity of \(B_C(\cdot)\), a sufficient condition for \(Q^C \geq Q^U\) is \(V_C(Q^U) \geq 0\) at \(Q = Q^U\). From Corollary 4, \(m^U(Q) \leq m^C(Q)\). In addition, as argued above, \(K^{(13)} = 0, K^{(3)} \leq 0,\) and \(K^{(23)} = K^{(33)} = 0\) if \(\rho(\cdot)\) has an affine structure.

\[
V_C(Q^U) = \mu'(Q^U) + (1 - \tau)\frac{K^{(3)}[y^C, m^C(Q^U)]}{1 - \beta q[m^C(Q^U)]}
\]

\[
= \mu'(Q^U) + (1 - \tau)\frac{K^{(3)}[y^U, m^U(Q^U)]}{1 - \beta q[m^U(Q^U)]}
\]

\[
\geq \mu'(Q^U) + (1 - \tau)\frac{K^{(3)}[y^U, m^U(Q^U)]}{1 - \beta q[m^U(Q^U)]} = 0.
\]

\(\square\)