Corporate Hedging:
What, Why and How?

by

Michael P. Ross

Haas School of Business
University of California, Berkeley

Spring 1996

Abstract

This paper explores the rationale for corporate risk management. Following Smith and Stulz (1985) and Mayers and Smith (1987), the assumption is made that firms can contractually commit to bondholders to maintain a particular risk management policy, or asset volatility. With that as a starting point, the essay derives the optimal hedge portfolio, examines this portfolio's robustness to variance-covariance misestimation, and proposes a new motive for corporate risk management; a firm that hedges its risk increases its optimal amount of debt and so realizes more tax benefits from leverage. Using the capital structure model of Leland (1994), three impacts of risk-reduction on shareholder value are measured: the increase in tax benefits, the reduction of bankruptcy costs and the reduction in the potential cost of the underinvestment problem. The essay's motivation is to serve as a guide to chief financial officers regarding the benefits of risk management and the sources of those benefits, so that risk management can be undertaken in a way that enhances shareholder value, rather than for its own sake.

The author wishes to thank Miguel Cantillo, Henry Cao, Sankar De, Aaron Edlin, Raul Espejel, Simon Gervais, Robert Goldstein, Avner Kalay, Hayne Leland, Stewart Mayhew, Richard Meese, Terry Odean, Mark Rubinstein, Eduardo Schwartz, Matt Spiegel, Brett Trueman, Hong Yan and seminar participants at the European Finance Association Symposium on Corporate Risk Management held in Vienna, the Conference on Corporate Risk Management in honor of Fischer Black held at UCLA and the Berkeley Program in Finance conference at Silverado for their numerous comments and helpful suggestions. Any errors remain my own.

1 Introduction

Financial economists have posited many theories as to why managers hedge. Motivated by Modigliani and Miller, whose seminal work points toward a lack of need for corporate hedging, academics have taken up the challenge of explaining this seemingly anomalous phenomenon. Equally anomalous is the managerial use of derivatives for purposes other than hedging. With well-publicized derivative fiascos plaguing numerous companies in recent years, a prescription for optimal derivative use may be in order.

In this paper, it is proposed that risk management enables firms to increase the optimal amount of leverage and so increase associated tax-benefits. The result of risk management may not be reduced risk for equityholders, since leverage increases. Rather, risk management enables the firm to substitute tax-benefitted risk, in the form of leverage, for non-tax-benefitted risk.

This paper will be organized as follows: In section 2, we will provide a brief overview of the hedging literature, offer some possible explanations for derivative use in general and introduce our principal explanation for hedging. In section 3, we will motivate and solve for the optimal hedge portfolio and illustrate the robustness of its risk-reduction to misestimation of the variance-covariance matrix of hedging instruments. We will also offer two testable propositions regarding how the character of hedging ought to differ among industries and how the derivatives boom has altered the nature of mergers and motivated spinoffs. In section 4, we demonstrate the value of risk-reduction to a firm's optimal capital structure and show that our particular leverage explanation of hedging generally provides more value for shareholders than the underinvestment leverage motive of hedging. In section 5, we discuss supporting evidence for our explanation for corporate hedging. In section 6, we conclude.

2 Explanations for Corporate Hedging and Derivative Use

Certainly, the most prolific researchers in the area of corporate hedging have been Clifford Smith, David Mayers and Rene Stulz. Although work on hedging individual projects dates back to the 1970s, Mayers and Smith published the earliest work on hedging corporations in their 1982 *Journal of Business* article, "On the Corporate Demand for Insurance." They suggest seven possible explanations for why corporations would insure their assets, even as their shareholders are likely diversified. With derivatives yet to undergo their awesome growth, the focus was on property and liability insurance, rather than on derivatives. Yet most of their insights were without loss of generality. Their four explanations that most shed light on derivative use, which were later elucidated by Smith and Stulz (1985), can be paraphrased as follows:

- Non-diversifiable stakeholders (employees, customers and suppliers) will demand expensive terms in contracts with a risky firm. For a variation on this theme, see Mayers and Smith (1990). Tufano (1995) offers strong evidence for managerial tenure and stock ownership as motives for hedging.

- The probability of a costly bankruptcy can be reduced.
- Tax Motives:
  - Limited or delayed deductibility of large losses—due to time-limits on loss carry-backs and carry-forwards and due to the government's abstention from participating in the firm's losses—make losses more costly than are gains beneficial. These effects are magnified for corporations currently already enjoying loss carryforwards or investment tax credits, as their ability to benefit from losses is further restricted.
  - Progressivity of corporate tax rates induces firms to smooth their profits. For any corporation large enough to use derivatives, the tax schedule is flat, as illustrated in the table on tax-rates.

A variety of other articles and authors have also provided explanations for corporate hedging:

- Leverage Stories:
  - Elimination of Risky Negative NPV Projects (Jensen & Meckling (1976)): Required hedging in bond indentures can reduce the firm's incentive to accept risky, negative NPV projects after debt issuance—and allow bondholders not to demand a corresponding discount on the bonds.
  - Acceptance of Positive NPV Projects (Myers (1977) and Mayers and Smith (1987)): A levered firm will refrain from some positive NPV projects when much of that value goes toward reducing the riskiness of debt outstanding. A bond covenant requiring hedging can help salvage the desirability to shareholders of some of those inefficiently rejected projects and reduce the cost of debt.
- Asymmetric Information
  - Discerning Managerial Ability (DeMarzo and Duffie (1992, 1995)): Suppose managers are better than shareholders at removing noise (via hedging) from corporate performance—due to superior information. Then by being induced to hedge, through a combination of risk-aversion, accounting mechanisms and compensation plans, shareholders can better discern managerial ability from corporate performance and exercise their option to fire incompetent managers, who are presumed to be ignorant of their own skill-level before hedging. In a 1995 revision of their 1992 working paper, shareholders can also better discern project quality and cancel bad projects.

  - Costly External Financing (Froot, Scharfstein and Stein (1993)): Suppose external financing is costly to current shareholders because potential investors are less well-informed than management. Hedging can reduce the probability that the firm will have insufficient internal funds to finance positive NPV projects that may arise and avert the need to issue securities at a discount to obtain financing. A brief counterargument would point out that if future positive NPV projects are large, then hedging may eliminate any possibility of avoiding external financing.

- Non-hedging derivative stories

  - Moral Hazard: Shareholders can expropriate bondholders by using derivatives to *increase* asset volatility.

  - Borrowing "Arbitrage": When borrowers face different spread differentials between the fixed and floating credit markets or between credit markets of different currencies, they can profitably engage in a swap if the type of debt they wish to issue is mutually that of their counterparty's comparative advantage.

We would like to suggest another explanation for corporate hedging. We believe that one of the strongest motivations for corporate hedging is the opportunity it provides for increased leverage and the tax-benefits that follow. This explanation is distinct from that of Mayers and Smith, in which they refer to reduced bankruptcy costs. As it turns out, bankruptcy costs do not necessarily decrease as the firm hedges because its chosen leverage increases. Indeed, most of the benefits to risk-reduction for an optimally-levered firm stem not from reduced bankruptcy cost[1], but from increased leverage and the resulting tax-benefits. Our explanation is also distinct from the underinvestment problem,[2] which neither focuses on maximization of optimal leverage nor on debt's tax benefits. However, as with the underinvestment problem, our prescription necessitates bond covenants that contain mandates for hedging.

Recent evidence from Hentschel and Kothari (1995) shows that leverage is highly positively correlated with derivative use. Dolde (1993) reports that the vast majority of companies surveyed use derivatives to reduce risk.[3] Putting these two findings together offers strong empirical support for our explanation. The fact that Hentschel and Kothari show derivative use to be only marginally related to equity volatility is consistant with our increased-leverage explanation[4] and poses a challenge to several competing explanations—perhaps most cogently to that of Froot, Scharfstein and

[1]For a statically-levered firm, most benefits of hedging do come from reduced bankruptcy costs, but these benefits are small realtive to the tax benefit afforded by increased optimal leverage.

[2]See Myers (1977) for an exposition on the underinvestment problem and Mayers and Smith (1987) for a hedging prescription for reducing its cost.

[3]As Smith (1995) points out, there may be a bias in this result; managers using derivatives to speculate may have elected to not respond to the survey.

[4]Note that much of the volatility-reduction achieved by hedging is offset by increased leverage.

Stein (1993).[5] Evidence by Tufano (1995) fails to support Froot, *et al* while offering weakly significant support for the leverage explanation, though his study focuses exclusively on the gold mining industry. A working paper by Géczy, Minton and Schrand (1995) offers strong support for Froot *et al*, while offerring weak support for our hedging explanation[6]

In this paper, the capital structure model of Leland (1994) will be employed to measure the benefits of risk management. In this model, the firm is levered with perpetual debt. All coupon payments to bondholders are tax-deductible, creating a tax benefit to leverage. On the other hand, a fixed proportion of firm value is dissipated upon default, creating a cost to leverage. This cost comes in the form of compensating bondholders with a higher coupon for the fact that a portion of firm value is lost coincident with any assumption of firm control by them. Trading off between these two effects, tax benefits and bankruptcy costs, produces an optimal capital structure. Bankruptcy is determined endogenously by the firm's shareholders as the point where keeping alive their call option on firm assets is no longer worth the cost of debt service. When the firm makes its leverage decision, it wishes to maximize the total value of debt plus equity, as any proceeds from debt issuance are used to buy back equity. This maximization is the same as maximizing tax benefits less bankruptcy costs over the amount of leverage.

As will be later demonstrated, hedging, by increasing the optimal amount of leverage, can increase the value of shareholders by 10⇔15% under very mild conditions. It will also be shown that hedging is a second-order effect for shareholder value. Most of the value that a CFO can generate for shareholders stems from leveraging in general, without first reducing firm volatility.

Proposition 1 *The value of the optimal capital structure is not necessarily monotonically decreasing in firm volatility.*

Proof: See appendix A.

Intuition: The bad-state payoffs of a low-volatility firm are higher than those of an equally-valued high-volatility counterpart. If default occurs only in the bad state and if bankruptcy costs are proportional, then the low-volatility firm will suffer a greater bankruptcy cost, as its firm value is greater in the bad state. This may make leverage less attractive than for the low-volatility firm than for the high-volatility firm.

Corollary 1 *The value of the optimal capital structure is monotonically decreasing in firm volatility when the value of the firm's assets follows a geometric Brownian motion.*

[5]Increased leverage increases the chance that a firm will need to seek external financing, should an attractive investment opportunity arise. A theoretical counterpoint to Froot *et al* is the prospect that future positive NPV projects may be large. If this is the case, then speculating is more likely than hedging to avert the need for external financing.

[6]As Froot *et al* discuss, Block and Gallagher (1986), Wall and Pringle (1989) and Nance, Smith and Smithson (1993) also weakly support our theory.

Proof: See Leland (1994) and Leland and Toft (1996).

If we may assume, as is common both in the literature and in industry, that the value of the firm's underlying assets is well-approximated by a geometric Brownian motion, then it follows that the goal of a value-maximizing manager is to minimize the volatility of his firm.[7] Volatility minimization is consistant with all of the hedging motives outlined above. It is, therefore, surprising that a formal approach to firm value risk minimization has never been proposed. Much work on optimal hedging has focused on how much foreign currency risk a US investor should hedge in a foreign stock portfolio.[8] With regard to optimal corporate hedging, Stulz (1984) calculates the firm's optimal hedge portfolio from the perspective of a risk averse manager, not the shareholder. Jacque's 1981 survey reviews literature on optimal foreign exchange hedges. But that work, as much that was to follow, focuses on hedging of individual positions or transactions and not on hedging the firm as a whole. FS&S suggest an optimal hedge portfolio that minimizes the volatility of "the shadow value of internal funds." One of the interesting implications of their work is that they show that options, and not futures alone, are generally required to hedge optimally, though their hedge is not on firm value. Hart and Ross (1994) introduce the continuous strike option, a straddle of which can perfectly hedge a firm facing multiple risks under certain specific circumstances.

3 The Optimal Hedge Portfolio

How shall we approach the problem of minimizing firm volatility? The idea that firm value is monotonically decreasing in asset volatility suggests that the ideal firm invests solely in the riskless asset. Of course such a firm cannot add any value for shareholders. A riskless firm could fully finance itself with debt and thus negate the taxes it would otherwise pay on the proceeds from its Treasury bill holdings. But such a firm would never be able to further distribute a cent to its equityholders, as all proceeds would go to bondholders who fully financed the t-bill purchase and bondholders would be no better off than had they purchased the t-bills directly.[9]

The purpose of a firm investing in risky assets—indeed, the purpose of a firm entirely—is to invest in real projects valued at a premium to their cost while reducing contracting costs and facilitating diversification and limited liability for shareholders. Firms have an ability to do this, that isn't arbitraged away via competition among firms, due to some propietary, perhaps patented, expertise and perhaps due to other barriers to entry. Evidence of this value-added can be seen in the high market-to-book multiple at which equities trade, though some of that premium is due to effects like depreciation and inflation. But even new firms trade at large multiples to book value

[7]By "firm volatility", we mean the volatility of the firm's assets, of which debt and equity are derivative securities.

[8]see Gardner and Stone (1995)

[9]Such a firm is simply a money market fund with no management fee.

that depreciation and inflation cannot explain.

Presuming, then, that the risky real assets in place are positive NPV investments[10], a firm's goal ought to be to minimize the firm's volatility subject to the constraint that it not liquidate its real assets. The firm's optimization problem can now be described as follows (for a derivation, please see appendix B):[11]

$$(\sigma^* V)^2 \;=\; \min_{\mathcal{D}} \; \mathcal{D}' \, \Omega \, \mathcal{D} \qquad \text{s.t. } \; R\,\mathcal{D} = V,$$

$$\text{where} \quad R \equiv [1 \; \underline{0}'].$$

$$\Rightarrow \mathcal{D}^* \;=\; V \cdot \frac{\lambda}{|\Omega_{-1,1}|}$$

$$\text{where} \quad \Omega \equiv \begin{bmatrix} \sigma_1^2 & \cdot\;\cdot\;\cdot & \sigma_{n+1,1}\Delta_{n+1} \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ \sigma_{1,n}\Delta_n & \cdot\;\cdot\;\cdot & \sigma_{n,n+1}\Delta_n\Delta_{n+1} \\ \sigma_{1,n+1}\Delta_{n+1} & \cdot\;\cdot\;\cdot & \sigma_{n+1}^2\Delta_{n+1}^2 \end{bmatrix}$$

$$\text{and} \quad \lambda \equiv \begin{bmatrix} +|\Omega_{-1,1}| \\ \Leftrightarrow|\Omega_{-1,2}| \\ \vdots \\ (\Leftrightarrow 1)^{j+1}|\Omega_{-1,j}| \\ \vdots \\ (\Leftrightarrow 1)^{n+2}|\Omega_{-1,n+1}| \end{bmatrix} \;=\; C_1(\Omega)$$

where:

- V is the value of the firm's underlying assets or the value of the all-equity firm
- Ω is the Variance-Covariance matrix of the firm's assets and n derivatives or hedge instruments
- $\sigma_{i,j}$ is the covariance between underlying assets i and j
- $X_{-i,j}$ is matrix X, with the 1^{st} row and the j^{th} column removed

[10] Assume that real assets in place are worth more unliquidated than liquidated even after accounting for the cost of their risk in terms of foregone leverage opportunities.

[11] Note that, prior to derivative use, $\mathcal{D} = [V \; \underline{0}']'$.

- $|X|$ is the determinant of matrix X or the absolute value of scalar X
- $\mathcal{D}$ is the vector of notional amounts of exposure, in dollars, to the $n+1$ underlying assets
- $\underline{0}$ is a length-n vector of zeros
- Δ_j is the delta of derivative asset j with respect to its own underlying.

Other notation:

- σ_1 is the volatility of the firm's underlying assets or, alternatively, of the unhedged, unlevered firm's equity
- σ_V is the volatility of the firm assets cum hedge portfolio
- z is the risk-reduction achieved through hedging, $z = 1 \Leftrightarrow \frac{\sigma_V}{\sigma_1}$
- Any $*$ superscript denotes a variable's value when optimized for either risk-reduction or leverage. For example, σ^* is the firm's asset cum hedge volatility when it is hedged with hedge portfolio $\mathcal{D}^*$ while C^* is the optimal coupon for a firm with volatility σ.
- $|\rho|$ is the determinant of the correlation matrix of the firm's assets and n hedge instruments.
- X_i is the i^{th} column of matrix X
- $X_{i,j}$ is the element of matrix X located in the i^{th} row and the j^{th} column
- $X_{i,-j}$ is the i^{th} column of matrix X with the j^{th} element removed
- $C(X)$ is the matrix cofactors of matrix X
- G_i is the i^{th} element of vector G
- G_{-i} is vector G with the i^{th} element removed.

3.1 Risk-Reduction

We shall now solve for the percentage by which the firm can reduce its asset volatility by hedging.

$$\begin{aligned}
\sigma^* &= \frac{1}{V}\sqrt{\mathcal{D}^{*\prime}\,\Omega\,\mathcal{D}^*} = \frac{\sqrt{\lambda'\,\Omega\,\lambda}}{|\Omega_{-1,1}|} \\
&= \sigma_1\sqrt{\frac{|\rho|}{|\rho_{-1,1}|}} \equiv \sigma_1(1 \Leftrightarrow z) \equiv \sigma_V \ < \ \sigma_1
\end{aligned}$$

where ρ is the correlation matrix of the firm and n hedge instruments and z is the percentage risk-reduction achieved.

Hedging With One Derivative:

$$\sigma^* \;=\; \sigma_1 \sqrt{1 \Leftrightarrow \rho^2} \;=\; \sigma_1 (1 \Leftrightarrow z)$$

Proposition 2 *z is independent of firm and hedging instrument volatilities.*

Proof: Simply note that $z = 1 \Leftrightarrow \sqrt{\frac{|\rho|}{|\rho_{-1,1}|}}$, which is independent of firm and hedge instrument volatilities.

Corollary 2 *The sign of any asset's set of cross-correlations is incidental to z.*

Proof: The sign of asset j's cross-correlations may be switched by multiplying column j and row j of ρ by $\Leftrightarrow 1$. This leaves the determinant of ρ unchanged[12].

3.2 A Sample Implementation

Now let us examine how much risk-reduction can be achieved using some of the most common hedge instruments. Given the S&P 500 as our hypothetical firm to be hedged, we will solve for the optimal hedge portfolio using the following derivatives.[13] Our hypothetical firm will have \$100 of underlying assets.

Using a hedge portfolio that includes the 10-year US treasury, the FTSE index, the three-month West-Texas Intermediate future, ECU and yen spot, our algorithm recommended dollar weights on these hedge instruments that were extremely reasonable in both direction and magnitude (see table (3.2)). With the exception of ECU, all hedge instrument weights have signs opposite their S&P correlations and with a magnitude of an identical order statistic to their S&P correlations. ECU is an exception due to its moderate colinearity with the yen. Cross-correlations can be seen in table (3.2). Despite correlations of a similar magnitude, the bond received a larger weight than the FTSE due to the bond's lower volatility and due to the FTSE's higher cross-correlations with the other hedge instruments. Hedge instrument volatilities can be seen in table (3.2).

The optimal hedge portfolio resulted in a z-factor of 16.8% for our hypothetical firm, enabling it to reduce its 8.59% volatility[14] down to 7.15%. Of course, this is

[12] Since we have twice multiplied $\rho_{j,j}$ by -1, $\rho_{j,j}$ remains one, as it must.

[13] Although the S&P 500 is composed of leveraged assets, this leverage does not locally affect correlations, upon which the risk-reduction factor solely depends.

[14] This is close to the lowest historical volatility ever observed on the S&P 500. Volatility data has been provided by RiskMetrics of JP Morgan. All σ's have been divided by 1.65–reversing RiskMetrics' significance adjustment. RiskMetrics uses a decay rate of 6% daily in calculating its variance-covariance matrix. Although this decay rate is good for value-at-risk purposes, it is far too rapid for effective hedging. Bob Litterman of Goldman Sachs has informed me that decay rates of about $\frac{1}{3}$% daily are near reasonable for optimal hedging purposes.

Asset	$\rho_{S\&P,j}$	$\mathcal{D}^*$
BOND	-.4403	+40.2%
FTSE	+.4094	-25.0%
OIL	-.1396	+4.7%
ECU	-.1257	-5.3%
YEN	-.1483	+4.9%

Table 1: Optimal hedge portfolio for the S&P 500 as a hypothetical firm on September 9, 1995.

$\rho_{i,j}$	S&P	BOND	FTSE	OIL	ECU
BOND	-.4403				
FTSE	.4094	-.2116			
OIL	-.1396	.0936	-.0569		
ECU	-.1257	.0732	-.3031	.0980	
YEN	-.1483	.0709	-.2704	.1036	.7473

Table 2: Correlations of hedge instruments and the S&P 500 on September 9, 1995. Data is due to RiskMetrics.

the optimal hedge portfolio, given the five hedge instruments employed. Certainly, inclusion of other hedge instruments would have enhanced our performance in the same way that adding explanatory variables to a regression increases the R^2.

A number of facts regarding our sample implementation are worthy of note. Assets having low absolute correlations with the firm contribute little to risk-reduction. For example, had we hedged with only the t-note and FTSE, our z-factor would have been 16.2%. Oil, ECU and yen combined reduced risk by only another 0.6%. In fact, if we were able to find a single hedge instrument with an absolute correlation with the firm of 0.555, we would have achieved the same amount of risk-reduction as did all five hedge instruments. A lone hedge instrument for which $|\rho_{1,j}| = 0.75$ achieves 33.9%

Asset	σ_j
S&P	.0859
BOND	.0774
FTSE	.1124
OIL	.1559
ECU	.1002
YEN	.1258

Table 3: Standard deviations of the S&P 500 and hedge instruments on September 9, 1995.

risk-reduction. This leads us to the following interesting implication. Firms in many industries can achieve far more risk-reduction by shorting their industry peers than by using currency, interest rate and commodity derivatives. Additionally, intra-industry hedges are more likely than traditional hedge instruments to have correlations with the firm that remain steady through time, thus giving them more potential as static hedges.

Unfortunately, the feasibility of this tactic is limited. Whereas there are no logistical supply constraints on derivatives, there are severe limits to the availability of shortable industry peers. Certain stocks may not be marginable at all by SEC regulations. A firm that constitutes a large percentage of its industry's market capitalization may be unable to short enough of its competitors' stock to hedge itself to the point that derivatives offer little additional hedge. Additionally, shorting stock is extremely costly, costing up to 3% per year even for institutions. Firms may wish to synthetically short their own stock[15] but such action would send a bad signal, which may make synthetic self-shorting prohibitively costly.[16]

Directly or synthetically self-shorting also has an interesting feedback property. Shorting 100% of one's own stock reduces risk by merely 50%.[17] It is necessary to short an infinite number of one's own shares to eliminate all risk. Likewise, there is a feedback effect when firms in the same industry short each other.

3.3 Robustness

A major concern in managing a hedge portfolio is whether it can achieve substantial risk-reduction even when the volatilities and correlations of hedge instruments cannot be forecasted accurately. It turns out that our optimal hedge portfolio is quite robust to variance-covariance misestimation. In fact, if all volatility estimates are off by the same percentage, then there is no loss of optimal risk reduction. With regard to misestimation of the correlation, the error's impact on risk-reduction is related to the square of the percentage misestimation, which mitigates errors of less than 100%.

Consider a single hedge instrument that has a true correlation with the firm of ρ and an estimated correlation of $\hat{\rho} = \rho(1 + \varepsilon)$. Then $\sigma_V = \sigma_1 \sqrt{1 \Leftrightarrow \rho^2(1 \Leftrightarrow \varepsilon^2)} = \sigma_1 \, (1 \Leftrightarrow z)$.

The top of figure (1) illustrates the relationship between correlation misestimation ε and the percentage of optimal risk-reduction achieved, $p \equiv \frac{z}{z^*}$. Notice that p is symmetric about zero, i.e. it is independent of the sign of ε. Also notice that for a range of true correlations, even a 50% estimation error can sustain three-quarters of

[15]Synthetic self-shorting can be achieved either by writing a call and buying a put, or by writing a forward, on one's own stock. Outright self-shorting does not really constitute a hedge, as it is equivalent to padding the firm's risky assets with non-risky cash, resulting in a firm with assets that are, on average less risky, despite being exposed to equal aggregate risk as before self-shorting.

[16]Equally, shorting a competitor can send a good signal and carry an added benefit.

[17]If s is the percentage of the firm that is synthetically shorted, then the percentage of risk reduction achieved is $\frac{s}{1+s}$, since the risk is shared between inside stockholders and outside synthetic stockholders in the proportion $\frac{1}{1+s}$ and $\frac{s}{1+s}$, respectively.

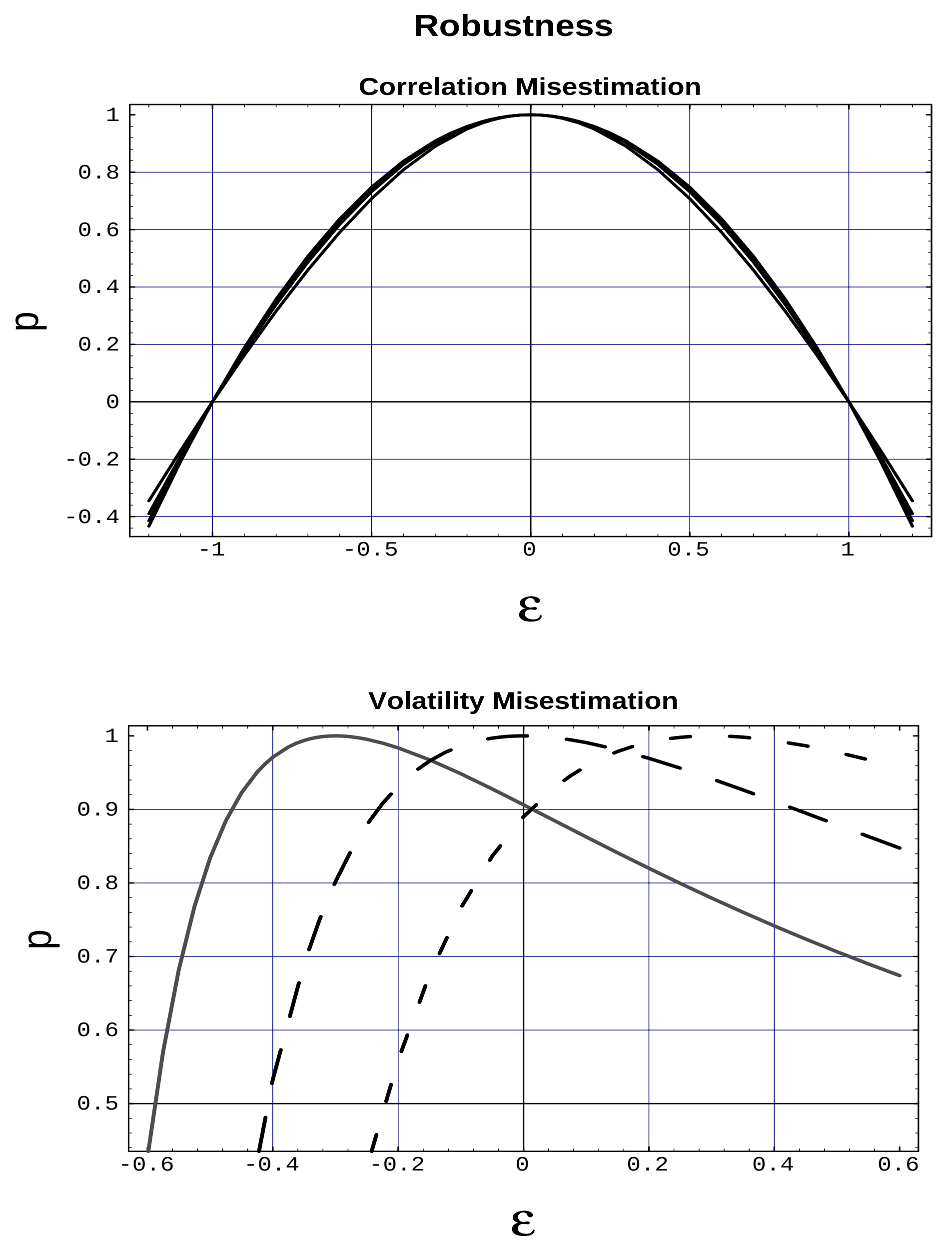


Figure 1: Above, we show robustness when there is one hedge instrument. In the top figure, we show the percentage of optimal risk-reduction achieved, p, for percentage correlation misestimates ε ranging from ⇔120% to +120%. Note that p is virtually invariant to our plots of .2, .4, .555 and .75 for the true correlation. In the bottom figure, hedge-instrument volatility is misestimated by percentage ε for firm volatility percentage misestimates of ⇔30%, 0% and +30% for the solid, long-dashed and short-dashed lines, respectively.

optimal risk-reduction. For p to be negative, it is necessary that $|\hat{\rho}| > 2 \cdot |\rho|$ or that the sign of $\hat{\rho}$ be opposite that of the true ρ. Thus, severe misestimation is necessary to significantly reduce or reverse the risk reduction benefits offered by our prescription.

The bottom of figure (1) shows the effect of misestimating hedge instrument volatility on p for three different firm volatility misestimates. Note that when the errors are the same ($\varepsilon_1 = \varepsilon_2$) there is no loss of optimal risk-reduction. This is a very convenient feature of our model, due to the tendency of volatilities in different asset classes to move up or down coincident with one another. This feature generalizes to an arbitrary number of hedge instruments. Unlike with correlation misestimation, p is no longer independent of the sign of the misestimate, appearing to be more sensitive to underestimates. This implies that risk managers may not wish to use their best unbiased estimate of hedge instrument volatilities in constructing their hedge portfolio, but may wish to purposely overestimate such volatilities. This leads to the following interesting result.

Hypothesis 1 *Value-maximizing firms may purposely underhedge, relative to the optimal hedge implied by the estimated hedge instrument volatilities, when complete and perfect static hedging is unavailable.*[18]

Reasoning: By corollary (6), $\mathcal{D}_j^*$ is inversely related to σ_j, implying:

$$\lim_{\hat{\sigma_j} \to 0} \hat{\mathcal{D}}_j^* = \infty \Rightarrow \sigma_V \to \infty \, \forall \, \varepsilon_j > 0.$$

$$\text{Likewise,} \qquad \lim_{\hat{\sigma_j} \to \infty} \hat{\mathcal{D}}_j^* = 0.$$

In the limit, severely underestimating hedge instrument j's volatility can result in an arbitrarily large weight on instrument j and thus an arbitrarily high volatility for the firm. Severe overestimation can do no worse than result in a zero weight on j.

This is not just a limiting result. Consider the case in which the volatility of hedge instrument $j \neq 1$ is misestimated by percentage ε_j. Then the we have:

$$\hat{\mathcal{D}}_j = (\Leftrightarrow 1)^{j+1} \cdot \frac{|\rho_{-1,j}|}{|\rho_{-1,1}|} \cdot \frac{\sigma_1 V}{\sigma_j (1 + \varepsilon_j)}$$

We would like to know if a volatility underestimate causes a larger deviation from the optimal weight $\mathcal{D}^*$ than an overestimate. The necessary and sufficient condition for this to hold is:

$$1 \Leftrightarrow \frac{1}{1 + |\varepsilon_j|} < \frac{1}{1 + |\varepsilon_j|} \Leftrightarrow 1 \Leftrightarrow \frac{2}{1 \Leftrightarrow \varepsilon_j^2} > 2$$

which obviously holds $\forall \, |\varepsilon| \leq 1$.

In our single hedge instrument example, when $\rho = .555$ and firm volatility is accurately estimated, overestimating hedge instrument volatility by 225% sustains half of optimal risk-reduction whereas an underestimate of 50% forfeits all hedge benefits.

[18] When static hedging is incomplete and/or imperfect (i.e. done with error), dynamic hedging via our prescription remains beneficial.

3.4 Derivatives: The Vanquisher of Conglomerates?

We would like to posit that, prior to the advent of derivatives, firms sought increased debt capacity via diversification. By merging with other firms that had a sufficiently low correlation, conglomerates could optimally borrow more than could the sum of their previously separate divisions.[19] The problem with this strategy is that the management of the parent company may have expertise that is limited to highly correlated businesses. Inefficiencies may well have resulted[20]

With the introduction of derivatives, the parent firm could spin off "non-core" divisions, i.e. divisions whose cash flows have a low correlation with the parent firm. Each entity could then hedge with derivatives, rather than through an inefficient conglomeration, and enjoy the benefits of increased debt. It is no mere coincidence, then, that the reversal of the conglomeration mania of the late sixties and seventies occurred in the eighties and nineties, just as derivatives became widely available. While the pace of mergers has not ebbed, they have recently tended to be between highly correlated firms, as with the recent spate of bank mergers.[21]

Hypothesis 2 *In light of this argument, it would be natural to posit that:*

- *mergers in the 1960s and 1970s were among firms of a lower correlation than those in the 1980s and 1990s.*
- *spunoff divisions in the 1980s and 1990s had a low correlation with their parent.*
- *conglomerate debt in the 1960s and 1970s exceeded amalgamated pre-merger debt.*
- *derivative use by the parent firm increased significantly after spinning off divisions.*
- *the market rewarded more, or punished less, acquiring firms of the 1960s and 1970s compared with acquiring firms of the 1980s and 1990s.*
- *the market rewarded spinoffs more in the 1980s and 1990s than in the 1960s and 1970s.*
- *amalgamated operating income suffered from merger activity in the 1960s and 1970s and benefited from merger activity in the 1980s and 1990s.*
- *amalgamated operating income benefited from spinoff activity in the 1980s and 1990s.*

[19]See Lewellen (1971) and Higgins and Schall (1975). Myers (1977) offers a counterargument.

[20]As Berger and Ofek (1995b) point out, Servaes (1995), Lang and Stulz (1994) and Berger and Ofek (1995a) have researched the cost of conglomeration during several periods for *already-existing* conglomerates.

[21]Kaplan and Weisbach (1992) find that takeover targets are more likely to be subsequently divested when the target's industry differs from the acquirer's.

4 The Value of Risk-Reduction

We would now like to assess the value of risk-reduction that stems from enhanced leverage opportunities. For this purpose, we shall make use of Leland's (1994) model of the firm's optimal capital structure. In that model, the value of the optimal leverage policy depends, among other variables, upon the volatility of the firm's underlying assets. In our implementation, the applicable volatility will be σ_V, the volatility of the firm's underlying assets cum hedge portfolio.

One limitation of Leland's model is that it solves for the firm's optimally-levered value given that interest rates are constant and that the firm issues exclusively infinite maturity debt. As he demonstrates, however, the presumption of constant interest rates has a small impact on bond values—on the order of 3%. With regard to debt being of infinite maturity, Leland and Toft (1996) demonstrate that the value of the optimally levered firm is monotonically increasing in the maturity of the debt issued, largely due to the lower endogenous bankruptcy level associated with longer term debt. Leland's model provides the value of the static, when-issued, optimal capital structure. Since the value of the dynamic optimal capital structure policy exceeds that of the static policy, Leland (1994) actually provides a lower bound on the value of the dynamic policy.[22]

In order to implement Leland's model, we will need estimates of the marginal tax benefits of debt and of bankruptcy costs. We shall also need to demonstrate that Miller's (1977) argument that, even in the presence of taxes, capital structure may not matter, does not apply in today's tax environment.

4.1 Why Miller (1977) Doesn't Apply

In his seminal presidential address to the American Finance Association, Merton Miller argues that when bankruptcy costs are small, traditional models that equate the marginal costs and benefits of debt would result in leverage far in excess of what we observe. Therefore, to justify the low leverage levels observed, there must be far smaller tax benefits to debt than corporate tax rates alone imply. The equilibrium Miller suggests is one in which companies in aggregate issue debt, at first to individuals or institutions with low marginal tax rates and then, as the low-tax clientele becomes satiated in debt holdings, to high marginal tax rate investors. So long as the marginal corporate tax rate exceeds that of the marginal corporate bond investor, the deductibility of interest is worth more to the corporation than the corporation must pay the investor in the form of a coupon sufficient to induce him toward debt and away from the less-taxed equity alternative. Eventually, so much debt gets issued that so many clientele of lower tax brackets become satiated in debt that the marginal

[22] I shall not here explore reasons that we do not consistantly observe firms issuing debt of as long a term as possible. I shall point out that the greater the asymmetric information over a firm's project quality, the more frequent the desirable monitoring, which shorter term debt facilitates. Also noteworthy is that Treasury Secretary Robert Rubin proposed in November, 1995, to eliminate the interest deductibility of corporate debt with a maturity exceeding 40 years.

debt investor's tax rate on bond income becomes high relative to his capped capital gains rate. He must then demand a coupon sufficiently high, to be willing to hold the debt, that all benefit to any individual firm from issuing debt is lost. This level of debt, which makes corporations indifferent between issuing more or less debt, is the equilibrium level of debt in the economy. No company benefits from its debt policy.

Using Miller's notation, τ_C is the marginal corporate tax rate, τ_{PS}^{α} is the marginal tax rate on equity returns for the marginal investor α and τ_{PB}^{α} is the marginal tax rate on ordinary/bond income for investor α.

In order for Miller (1977) to hold, there needs to be a marginal investor α for whom the marginal corporate benefit τ of issuing another dollar of debt equals zero.

$$\tau = 1 \Leftrightarrow \frac{(1 \Leftrightarrow \tau_C)(1 \Leftrightarrow \tau_{PS}^{\alpha})}{(1 \Leftrightarrow \tau_{PB}^{\alpha})} = 0$$

Let us now assess what values we ought to use for the various tax rates. We shall conservatively ignore the effect of state tax rates, which would tend to increase τ. Deductions, credits and carryforwards not withstanding, the τ_C is virtually flat at .35, as table (4.1) illustrates. A reasonable estimate for τ_{PB}^{α} might be 31%. Those in a higher tax bracket probably ought never to buy taxable debt. The rate is likely to be substantially lower, as well, for other reasons. First, bondholders have a free option to defer capital gains while realizing capital losses early.[23] Additionally, Litzenberger and Ramaswamy (1979) estimated τ_{PB}^{α} at .28, a far smaller percentage of the top marginal rate at that time than is .31 today.[24] Finally, assume τ_{PS}^{α} to be .20. Capital gains rates today are capped at .20. The effective rate is lower due to deferral value and due to forgiveness upon death but the effective rate on stock earnings is higher due to the dividend component of the return on equities. Taking these offsetting effects into account, we will use a tax rate on stock earnings of 20%. Together, these assumptions result in a marginal tax benefit for debt of .25. In fact, for Miller (1977) to apply, we would today need, ceteris parabis, $\tau_{PB}^{\alpha} = .48$, higher than today's maximal rate of .396.

As the table (4.1) illustrates, a τ near zero is not feasible for any set of rates in today's tax structure.

4.2 Estimating Bankruptcy Costs

Given that Miller (1977) does not apply in today's tax environment, it must be the case that bankruptcy costs are the reason that firms are not fully debt financed. For the same reason, it must be the case that bankruptcy costs are far greater than direct

[23] See Constantinides and Ingersoll (1984)

[24] Actually, Litzenberger and Ramaswamy estimated the implied tax rate of the marginal investor in dividend-paying stocks. This investor would arguably be the highest tax-rate investor in dividend-paying stocks. The lowest tax-rate investor for dividend paying stocks is, however, the marginal investor we seek, i.e. the highest-tax-rate investor in taxable bonds.

Marginal Corporate Tax Rates	
Taxable Income	τ_C
\$0—50K	.15
—\$75K	.25
—\$100K	.34
—\$335K	.39
—\$10M	.34
—\$15M	.35
—\$18.$\bar{3}$M	.38
\$18.$\bar{3}$M +	.35

Table 4: Marginal corporate tax rates, post-1992.

τ		τ_{PB}^{α}				
	$\tau_C = .35$	.15	.28	.31	.36	.396
τ_{PS}^{α}	.15	.35	.23	.20	.14	.09
	.20	.39	.28	.25*	.19	.14
	.28	.45	.35	.32	.27	.23

Table 5: Marginal tax benefit of debt, for different marginal investor profiles.

costs alone.[25] Direct costs include, primarily, legal expenses. Indirect costs include the cost of fire-sales on hard assets, the reticence of suppliers to sell to a firm on the verge of default, the tacit boycott of customers concerned with the firm's future capacity to honor warranties, supply spare parts or upgrade product, the bankruptcy trustee's inefficient administration of the firm and, not incidentally, management's time and attention consumed by bankruptcy proceedings.

An imperfect way to measure total bankruptcy costs (direct plus indirect) would be to observe recovery rates on defaulted debt. Subtracting the recovery rate from one would yield total percentage bankruptcy costs, but only if there exists a positive net worth covenant, which is difficult to enforce due to asymmetric information between lenders and shareholders.

Without a credible positive net worth covenant, total percentage bankruptcy costs are one less the ratio of the recovery rate to the default time's asset value as a fraction of principal owed. We could attempt to observe the total debt and equity value of a firm prior to news of default in hopes of measuring the percentage of debt's face value constituted by the firm's total asset value at the time of default. But expected bankruptcy costs would already be discounted into the near-default debt and equity prices.

[25] Warner (1977) measured direct bankruptcy costs for the railroad industry at 1.3% for the period 1933—1955.

Fortunately, Leland (1994) can help extricate us from this measurement quagmire. He solves for the endogenous bankruptcy level of the firm, at which the firm's shareholders are unable to raise proceeds from secondary equity issuance to service debt payments. We would intuit that firms tend to default at asset values below debt principal for several reasons. Such a policy lowers the probability of bankruptcy relative to a zero-net-worth default rule (positive net-worth bond covenant). It also lowers the cost of bankruptcy when it occurs, if bankruptcy costs are monotonically increasing in firm value. This policy also decreases the fraction of bond face value that is returned to bondholders in bankrupt states, when payments to bondholders are generally not tax-deductible. Finally, a negative net-worth default policy is consistant with the moral hazard of shareholders to avoid default whenever possible. Indeed, Leland's model is consistent with a negative net-worth default policy.

Let us apply Leland's work to solve for total percentage bankruptcy costs. We shall require data on recovery rates to proceed with measurement. Altman and Bencivenga (1995) provide such data from 1985—1994 for various classes of debt. Their data on unsecured senior debt is most applicable, as this is the class of debt for which Leland's model provides endogenous bankruptcy levels. Let us define the following:

- RR, recovery rates on unsecured senior debt
- b, fixed percentage bankruptcy costs
- V_B, the endogenous bankruptcy level
- F, the face value of debt outstanding
- β, the percentage of residual asset value paid to equity when strict priority is not respected.

$$
\begin{aligned}
RR &= (1 \Leftrightarrow b)(1 \Leftrightarrow \beta)\frac{V_B}{F} \\
\Rightarrow b &= 1 \Leftrightarrow \frac{RR}{\frac{V_B}{F}(1 \Leftrightarrow \beta)}
\end{aligned}
$$

Using parameters specified in the caption to figure (2) and a recovery rate of 50%, we obtain total bankruptcy costs of 22%. Our measurement is imperfect for several reasons. Leland (1994) is for infinite maturity debt. Applying Leland and Toft (1996) is better-advised. Our measurement accepts all of Leland's assumptions of constant volatility and interest rates and that of geometric Brownian motion of firm asset value. We would also do better to estimate b for each of the firms in Altman and Bencivenga's sample separately and then take a cap-weighted average. Although a careful empirical study of bankruptcy costs is beyond the scope of this paper, we are pleased to introduce the basis for a sound methodology for total bankruptcy cost estimation. With our ad-hoc estimate of 22% for bankruptcy costs, we are finally equipped to measure the value of risk-reduction.

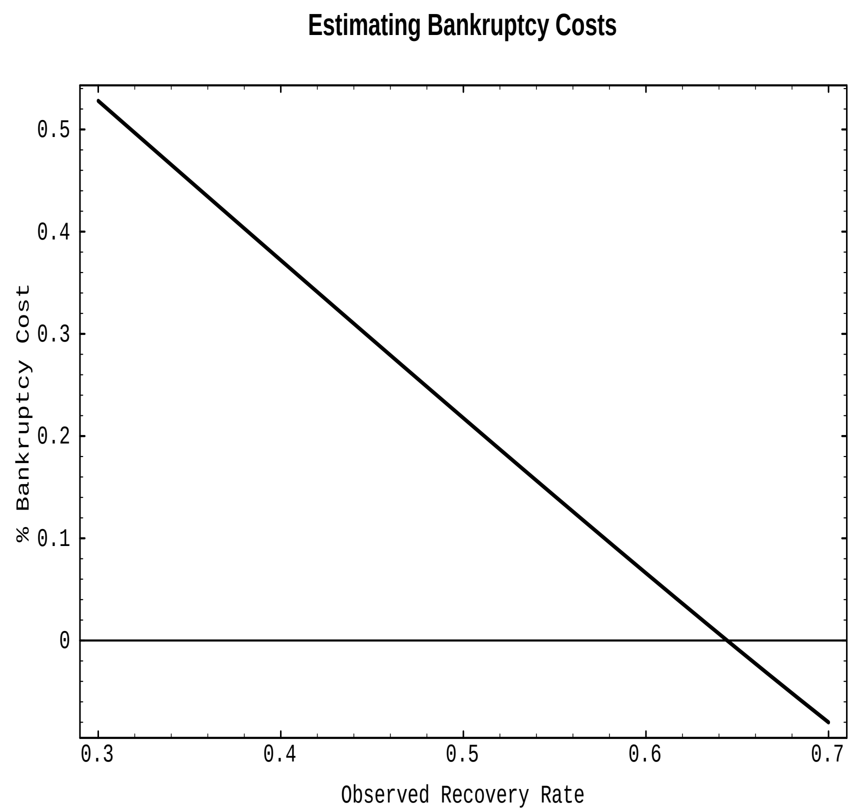


Figure 2: Assume: $\beta = 0, \sigma = .2, r = .09, \tau = .27$. Note that r and τ are higher than elsewhere in this paper since we are observing recovery rates from 1985-94. τ is higher because τ_{PB}^{α} was lower during that period relative to other tax rates than it is today. Recall that β is the share of the firm going to equityholders in bankruptcy.

4.3 Another Sample Implementation

Consider an unhedged firm that is optimally levered, given its current volatility. This firm currently has limited incentive to hedge, as doing so enhances bondholder value at the expense of shareholders. However, if the firm calls or otherwise repurchases the outstanding debt, it can optimally reissue a larger amount of debt, provided it can credibly commit to risk-reduction. Mayers and Smith (1987) recommend a debt covenant that obliges hedging as a means of mitigating the underinvestment problem. We propose that such a covenant can also facilitate increased shareholder value due to the greater tax benefits accruing to a more highly optimally levered firm.[26]

[26] When adherence to such a covenant is difficult to monitor or enforce, short term debt may dominate long term debt, as the reputational value of hedging for the sake of future debt issuance can then exceed the value of expropriating current bondholders. Perhaps we observe firms statically hedging individual positions rather than dynamically hedging asset value because the former activity

Suppose our sample firm is described by the following parameters:

$\underline{V}$	$\underline{r}$	$\underline{\sigma}$	$\underline{\tau}$	$\underline{BC}$
100	.067	.20	.25	.22

Then, by Leland (1994), the optimal capital structure consists of:

$$\begin{aligned} Equity &= \$28.54 \\ Debt &= \$90.47 \\ MktCap &= \$119.01 \end{aligned}$$

Suppose the firm recognizes that hedge instruments now exist capable of reducing its σ to .15. As long as firm assets have not performed too poorly since the outstanding debt was issued, the firm can well serve its shareholders by buying back the current debt. The buyback can be financed either with a bridge loan or with the new debt issue. Of course, the new debt issue must, either through an enforeable covenant or via reputational value, induce the firm to adhere to its 25% risk-reduction strategy. With $\sigma = .15$, the optimal capital structure will be:

$$\begin{aligned} Equity &= \$22.81 \\ Debt &= \$99.51 \\ MktCap &= \$122.32 \end{aligned}$$

- Step 1) Buy Back Debt with a Bridge Loan for $90.47
- Step 2) Reduce σ to .15
- Step 3) Issue $99.51 in New Debt[27]—Repay Bridge Loan
- Step 4) Distribute the $9.04 of Extra Debt Financing to Shareholders[28]
- Step 5) Equity Gains $\$9.04 + E_{after} - E_{before} = \$3.31$—an 11.6% Gain!

Percentage gains of this magnitude are common over a wide range of parameter values, as table (6) and figure (3) illustrate. Were the firm simply to reduce its risk by 25% without calling the old debt, it would forfeit $1.89 to bondholders, since levered equity is like a call option, the value of which increases in σ.

Proposition 3 *When shareholders are atomistic, callable debt dominates non-callable debt.*

is more easily monitored by bondholders.

[27]Cum hedge covenant or of a short enough maturity for reputational value to induce continued hedging.

[28]This may be done via dividend or stock buyback.

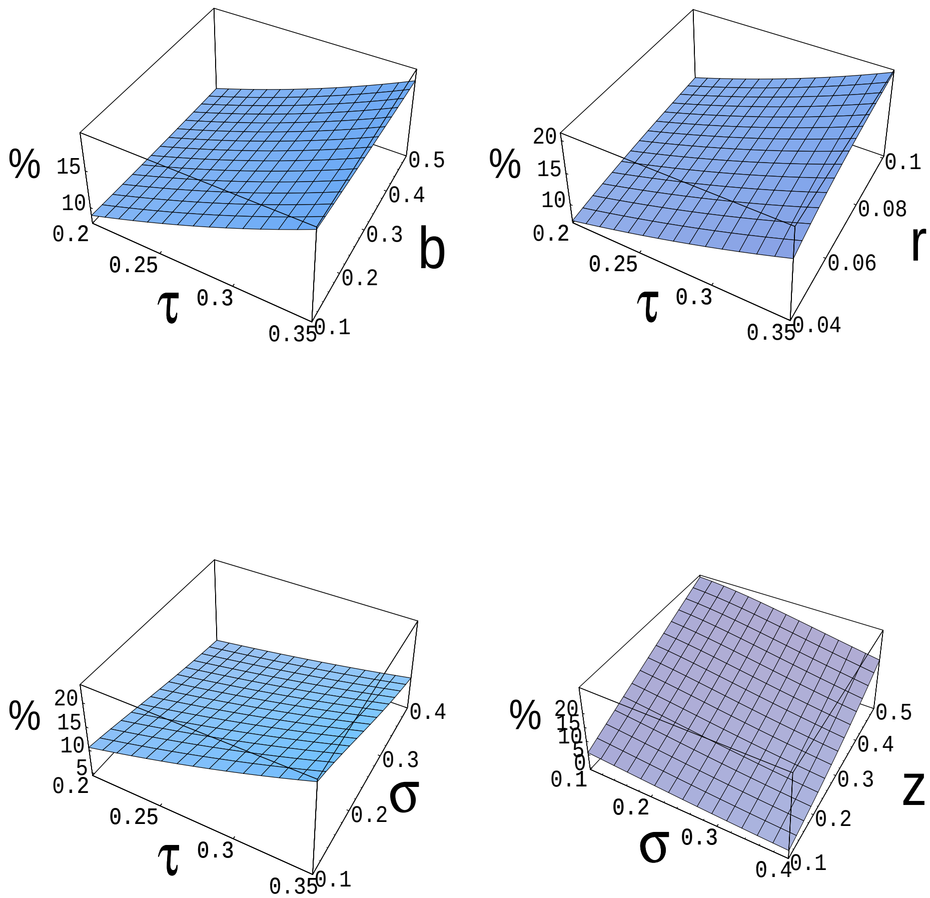


Figure 3: Unless an x-axis or y-axis variable, parameters are as follows: $V = 100$, $r = 0.067$, $\sigma_1 = 0.2$, $\tau = 0.25$, $b = 0.22$ and $z = 0.25$. The z-axis measures the percentage gains attributable to risk-reduction that accrue to equityholders of an optimally-levered firm . Risk-reduction is 25% for all but the bottom-right figure.

Proof: Although, in the example, the firm repurchased non-callable debt at its no-arbitrage value,[29] if bondholders are atomistic, then free-ridership may prevent the firm from realizing any benefit from our prescription.[30] Callable debt allows the firm to benefit from future expansion of the set of hedge instruments without having to issue very short term debt, which would decrease the value of optimal leverage.[31]

Two testable hypotheses follow from the foregoing analysis.

Hypothesis 3 *Callable debt has become increasingly predominant as the range of hedge instruments has increased.*

[29]Leland's model presumes the market to be dynamically complete with respect to the firm's debt and equity.

[30]See Grossman and Hart (1980)

[31]See Leland and Toft (1996).

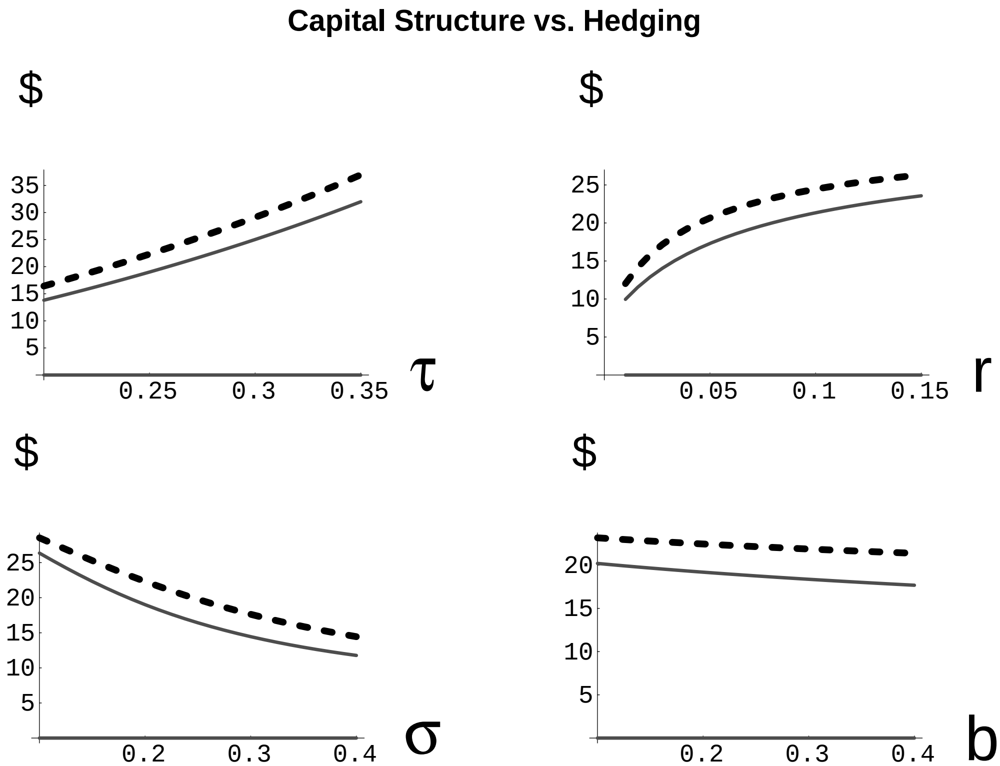


Figure 4: Unless an x-axis variable, parameters are as follows: $V = 100$, $r = 0.067$, $\sigma_1 = 0.2$, $\tau = 0.25$, $b = 0.22$ and $z = 0.25$. The solid line is the contribution to $100 unlevered firm value of optimal leverage when asset volatility is 0.2. The dotted line is the the contribution to $100 unlevered firm value of optimal leverage when asset volatility is 0.15. The difference is the value of 25% risk-reduction.

It is interesting to note that shareholders received $19.01 of benefit from the pre-hedge capital structure optimization and only $3.31 extra from risk-reduction. It is thusly apparent that the value of risk-reduction is truly a second order consideration relative to the pursuit of optimal capital structure. This observation is valid for a wide range of parameters, as figure (4) illustrates.

As tables (6) and (7) demonstrate, however, 25% risk-reduction can make as large a percentage contribution to equityholders as can the initial optimal leverage decision, since after initial levering and prior to risk-reduction, equity constitutes a smaller percentage of firm value. Therefore, the fewer dollars of benefit that risk-reduction confers can nevertheless confer a large percentage-gain to equityholders.

4.4 A Comparison of Leverage Stories of Hedging

Having demonstrated the value of hedging for the optimal capital structure, we are prepared to compare the value of this hedging motive against other leverage stories. In particular, it would be interesting to examine when risk-reduction contributes more value in enhancing the optimal capital structure and when it contributes more value

Value Added to Equity (%) from Risk-Reduction

		Risk-Reduction Factor, z				
		.1	.2	.3	.4	.5
σ_1	.10	5.7	11.4	17.1	22.7	28.1
	.15	5.0	10.3	15.8	21.5	27.2
	.20	4.3	9.1	14.3	19.8	25.7
	.25	3.7	7.9	12.6	18.0	23.9
	.30	3.1	6.8	11.1	16.1	21.9

Table 6: Parameter values are $r = 0.067$, $\tau = 0.25$ and $b = 0.22$. Presume that the firm is initially optimally levered for its initial volatility σ_1, buys back its outstanding debt, reduces volatility by factor z and optimally relevers. The firm uses a bridge loan to finance its initial debt repurchase and uses the proceeds of its second debt offerring both to repay the bridge loan and to buy back shares.

Value Added to Equity (%) from Leverage

σ_1	.10	26.4
	.15	22.3
	.20	19.0
	.25	16.4
	.30	14.4

Table 7: Parameter values are $r = 0.067$, $\tau = 0.25$ and $b = 0.22$. Presume that the firm is initially unlevered, issues an optimal amount of debt, and uses the proceeds to repurchase equity.

in mitigating the underinvestment problem.

First, we shall need to compute the potential value of risk-reduction in fixing the underinvestment problem.

Proposition 4 *The* maximal *cost of the underinvestment problem in a two-date setting is:*

$$e^{-rt} \int_0^{V_B} (V_B \Leftrightarrow V_s)\, f(V_s)\, dV_s, \qquad (1)$$

where $f(V_s)$ is the risk-neutral probability density function over state-s firm asset values.

Proof: The value described in (1) is the expected value to shareholders of the right to put the firm to the bondholders. If shareholders were to accept a free project that paid $(V_B \Leftrightarrow V_s)$ in all states $s: V_s \leq V_B$, then no value would accrue to shareholders. The social loss is then equal to (1), since the project is free. If the project additionally

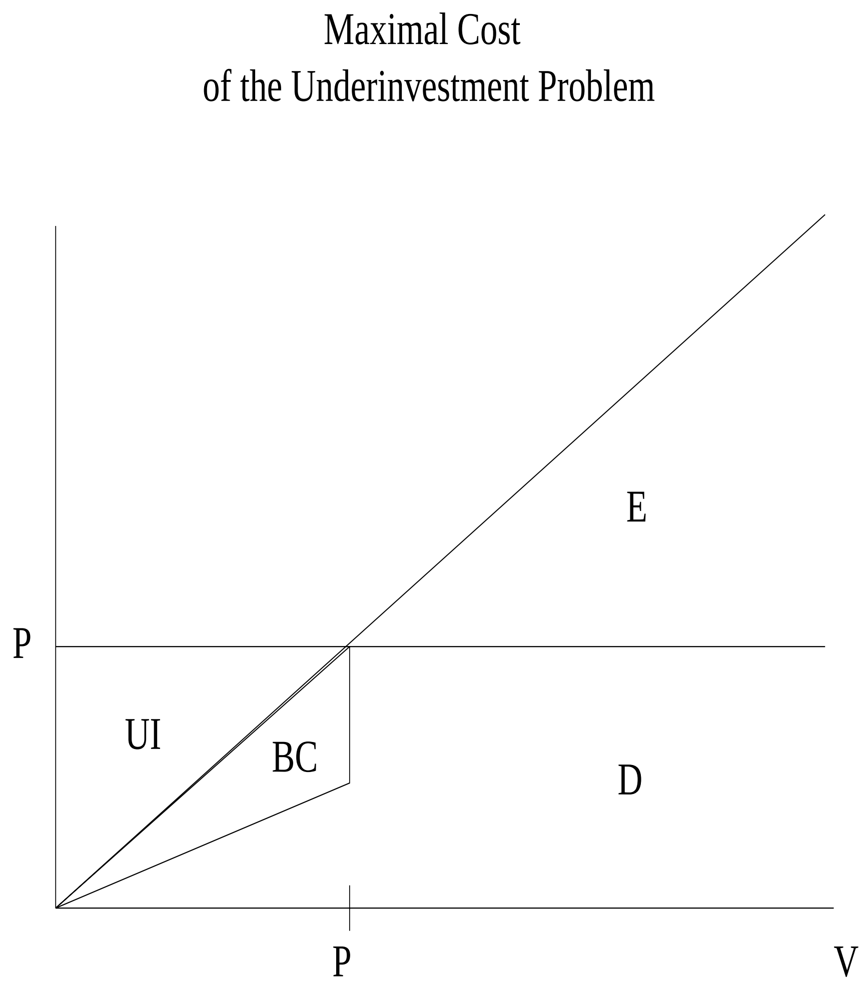


Figure 5: Area E represents the claims of equity, area D represents the claims of debt, area BC represents state-claims eaten by bankruptcy costs and UI represents the maximal amount of state-claims that can be lost to the underinvestment problem.

returned $\epsilon > 0$ in any state s, then shareholders would benefit from accepting the project and the social loss would be nil.[32]

Refer to figure (5). Proposition (4) states that the maximal cost of the underinvestment problem is represented by area UI. Consider a firm that is worth E + D. Suppose a project becomes available to management and that management is behaving in the best interest of equityholders. If the project offers payoffs in states and in amounts represented by area UI, then the project will be rejected, as it offers nothing to equityholders since the entire region below P on the y-axis has been promised to bondholders. If the project offers any *additional* positive NPV payoffs, such payoffs will accrue to equityholders and the project will be accepted. Since bondholders can

[32]If such projects were non-mutually exclusive and became repeatedly available, then the loss could exceed (1) if and only if management is myopic. With foresight of this repeated availability, management will accept such projects early on in order to benefit from those to follow.

assess, ex-ante, that the moral hazard of the underinvestment problem will deprive them of beneficial projects, they will accordingly pay less for the debt. In this way, the cost of the underinvestment problem falls on equityholders. If equityholders hedge then the probability density and, therefore, the value of area UI is reduced. If equityholders increase leverage after hedging, then, although the probability density over this area is reduced, the size of area UI generally expands. This can make the sign of the change in the value of area UI ambiguous in risk-reduction for a firm that levers optimally.

Corollary 3 *Callable debt mitigates the underinvestment problem, eliminating it entirely when call is not suboptimal.*

Proof: When debt is callable, the maximal social loss of underinvestment is zero, unless the call price of debt exceeds its uncalled value (given rejection of the positive NPV project) by more than (1), in which case the social loss is (1). Since this can only be less than or equal to (1), callable debt dominates non-callable.

Since the value of mitigating the underinvestment problem depends only indirectly on τ and bankruptcy costs, we might think that it should dominate enhanced leverage as a hedging motive when corporate taxes are low or when bankruptcy costs are high, i.e. when leverage is only slightly desirable in the first place. However, since under these circumstances, a firm is likely to have little debt, the cost of the underinvestment problem is likely to be small. Conversely, when τ is large and b small, firms are likely to be highly levered. The cost of the underinvestment problem is likely to be large as is the benefit of risk-reduction to increased leverage. It is not obvious which explanation will dominate, since both motives tend to be valuable or not valuable in similar circumstances. We can write down the formulae for the maximal value to each hedging motive by applying Leland (1994). Note, however, that whereas management has discretion over enjoying the maximal value of enhanced leverage, an investment project generating the maximal cost of the underinvestment problem is extremely unlikely to arise, since if it returned an infinitesimal amount more, it would generate no social cost whatsoever.

In Proposition (4), we described the maximal value of *eliminating* the underinvestment problem via complete hedging or via debt retirement. We shall now describe the maximal value of *reducing* the underinvestment problem through partial risk-reduction in the more general setting of Leland (1994).

The maximal social cost of the underinvestment problem is equal to the value of shareholders' limited liability, which is the same as the value of shareholders' right to default on the debt. This, in turn, equals the value of the firm's promise to bondholders, contingent on it being always fulfilled, less the value of the debt and less the present expected value of bankruptcy costs.

Let's consider a firm that is initially optimally levered. It calls its debt, and issues new debt with a promise to hedge. However, it only issues an amount of debt equal to what it previously had outstanding, i.e. it does not re-optimize. This firm is only concerned with the underinvestment problem and is not concerned with increased optimal capital structure vaule. Reoptimizing the leverage will only serve to increase the amount of debt outstanding and will thereby reduce the value of risk-reduction to mitigation of the underinvestment problem. This firm will enjoy some tax-benefits and some reduced bankruptcy costs from its risk-reduction even without re-optimizing. But we shall here only measure value accruing to this firm from reduction of the maximal underinvestment cost.

Additional notation:

- $C(\sigma)$, the coupon paid by a firm with volatility σ
- $D(C, \sigma)$, the debt value
- $x \equiv \frac{2r}{\sigma^2}$
- $UI^{(*)}(\sigma)$, the maximal cost of underinvestment to a(n) (optimally-levered) firm with volatility σ
- TB, the present expected value of the tax-benefit of debt
- BC, the present expected value of the bankruptcy cost of debt
- VA_{UI}, the maximal social benefit of mitigating the underinvestment problem through risk-reduction
- VA_{OCS}, the value-added to shareholders of using risk-reduction to increase the value of the optimal capital structure.

The maximal value of risk-reduction to mitigating the underinvestment problem: *when the firm* only initially *adheres to an optimal capital structure policy*[33], and leaves leverage unchanged after hedging, is:

$$\begin{aligned}
UI^*(\sigma_i) &= \frac{C^*(\sigma_i)}{r} \Leftrightarrow \left[D^*(\sigma_i) + bV_B^*(\sigma_i, C^*(\sigma_i))\left(\frac{V}{V_B^*(\sigma_i, C^*(\sigma_i))}\right)^{-x(\sigma_i)}\right] \forall\, i \\
UI(\sigma_v) &= \frac{C^*(\sigma_1)}{r} \Leftrightarrow \left[D^*(\sigma_V, C^*(\sigma_1)) + bV_B^*(\sigma_V, C^*(\sigma_1))\left(\frac{V}{V_B^*(\sigma_V, C^*(\sigma_1))}\right)^{-x(\sigma_V)}\right] \\
VA_{UI} &= UI^*(\sigma_1) \Leftrightarrow UI(\sigma_v)
\end{aligned}$$

when the firm calls the old debt, reduces risk and credibly binds itself to hedge into the future, and reissues an optimal amount of debt is:

$$VA_{UI}^* = UI^*(\sigma_1) \Leftrightarrow UI^*(\sigma_V)$$

[33]When there is no benefit to debt, mitigation of the underinvestment problem is uninteresting—simply eliminate all debt.

The value of risk-reduction in enhancing the value of the optimal capital structure is:

$$
\begin{aligned}
TB^*(\sigma_i) \Leftrightarrow BC^*(\sigma_i) &= \frac{\tau C^*(\sigma_i)}{r}\left[1 \Leftrightarrow \left(\frac{V}{V_B^*(\sigma_i,\, C^*(\sigma_i))}\right)^{-x(\sigma_i)}\right] \\
&\Leftrightarrow bV^{-x(\sigma_i)} V_B^*(\sigma_i,\, C^*(\sigma_i))^{x(\sigma_i)+1} \\
VA_{OCS} &= TB^*(\sigma_v) \Leftrightarrow BC^*(\sigma_v) \Leftrightarrow TB^*(\sigma_1) + BC^*(\sigma_1)
\end{aligned}
$$

In forthcoming graphs, we will demonstrate that the above formulae imply that our leverage story of hedging comes close to dominating the underinvestment story. We will also show that most value from increased optimal leverage stems from increased tax-benefits and not from reduced bankruptcy costs (see figure (6). Because there always exist parameters that enable either of the other leverage-based hedging stories to provide more value from risk-reduction than ours, the dominance of increased tax-benefits over reduced bankruptcy costs or decreased abstinance from positive NPV projects is not strict and only holds for most reasonable parameters. However, using Leland (1994), we may convey a closed-form sense of the dominance of enhanced capital structure value over mitigation of the underinvestment problem and bankruptcy costs.

Proposition 5 *The value of the optimal capital structure is monotonically increasing in risk-reduction for an optimally-levered firm while bankruptcy costs and the maximal cost of the underinvestment problem are not.*[34]

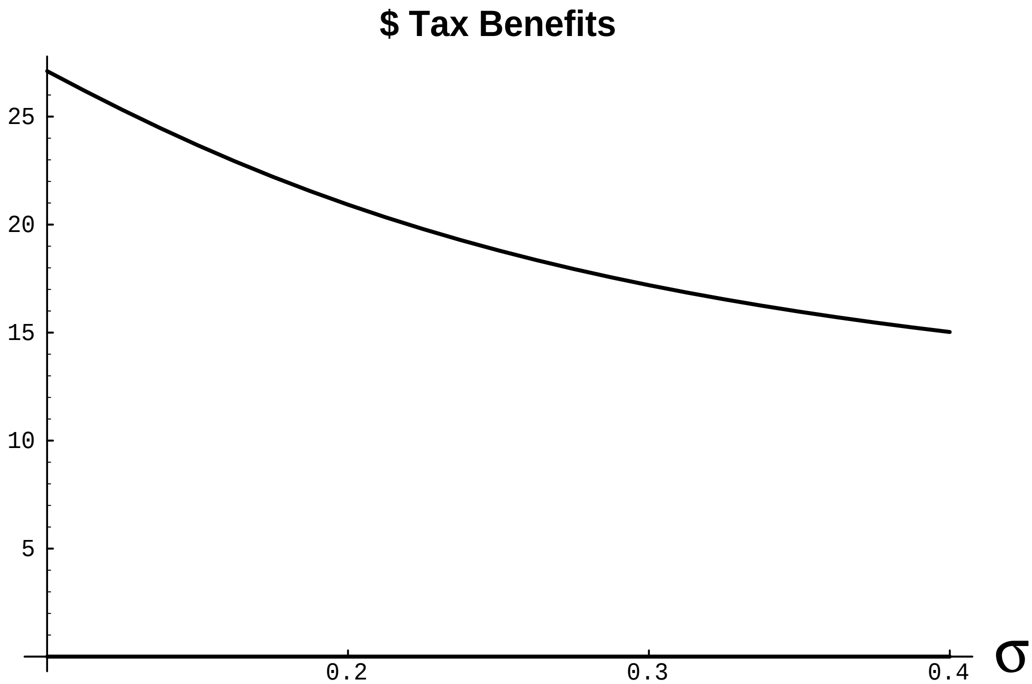
Proof: For monotonicity of the value of the optimal capital structure in risk-reduction, see corollary (1). Using Leland (1994), it can be shown that:

$$
\begin{aligned}
&\frac{\partial UI^*}{\partial \sigma_v} < 0\,\forall\,\theta \ \text{ iff } \frac{\partial\left\{(1 + \frac{1}{x})\log[1 + xf]\right\}}{\partial x} > \frac{\tau f}{(1 \Leftrightarrow b) + \tau x f}\ \forall\theta \\
&\frac{\partial BC^*}{\partial \sigma_v} < 0\,\forall\,\theta \ \text{ iff } \frac{\partial\left\{(1 + \frac{1}{x})\log[1 + xf]\right\}}{\partial x} > 0\,\forall\theta \\
\text{where}\quad & f \equiv 1 + \frac{b(1 \Leftrightarrow \tau)}{\tau} \\
\text{and}\quad & \theta \text{ is the vector of all relevant parameters.}
\end{aligned}
$$

It is easily shown that neither condition holds for τ small.

The intuition for the non-monotonicity of bankruptcy costs and the maximal cost of the underinvestment problem in firm volatility is clear. As the firm reduces risk, it increases leverage. The net result is a reduction in the probability of bankruptcy, despite higher debt levels, as Leland (1994) shows. However, although bankruptcy occurs less frequently, it is more costly when it does occur, due to the greater outstanding debt. The product of this greater cost with the lower probability is ambiguous.

[34] Using the assumptions of Leland (1994).



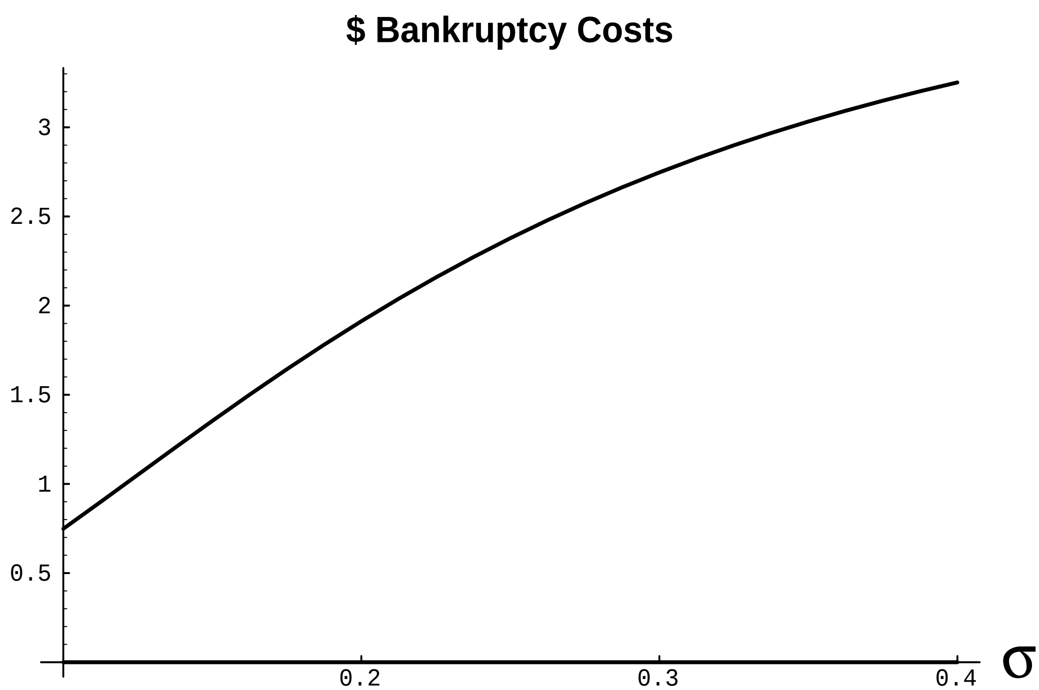


Figure 6: Once again, $V = \$100$, $r = 0.067$, $\tau = 0.25$ and $b = 0.22$. Notice that as we reduce σ from .4 to .2 and, then, from .2 to .1, about \$5 in tax benefits accrue to equityholders but only about \$1.25 in bankruptcy cost reduction.

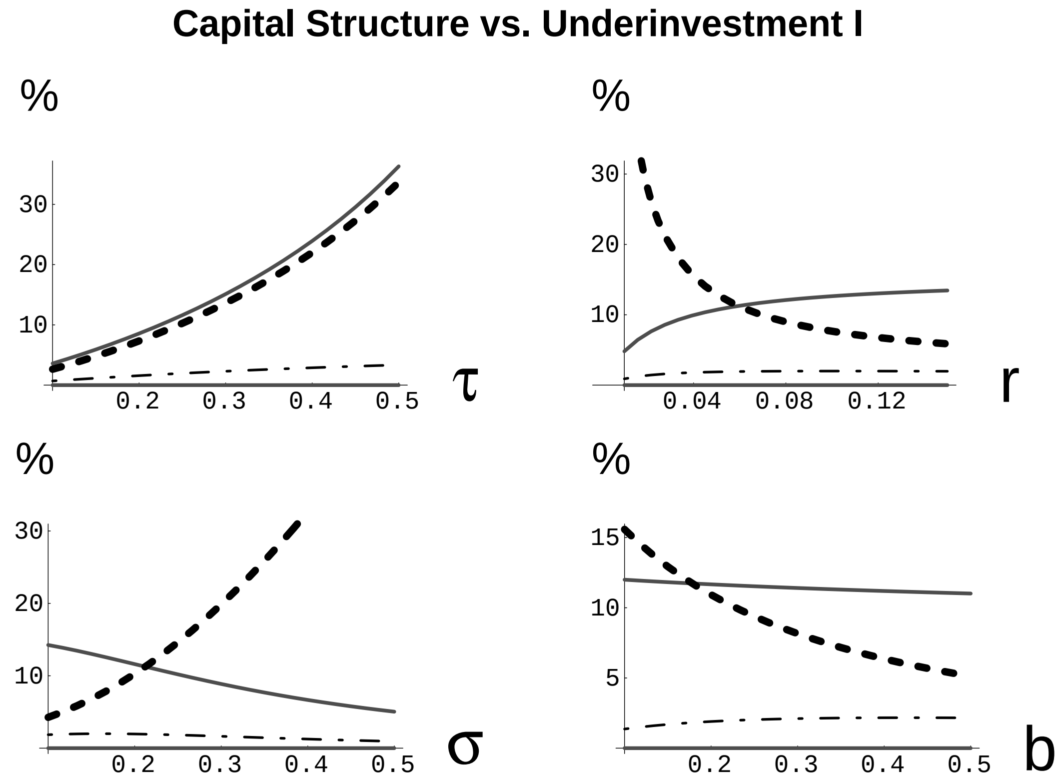


Figure 7: As usual, we have $V = \$100$, $r = 0.067$, $\sigma_1 = 0.2$, $\tau = 0.25$, $b = 0.22$ and $z = 0.25$. The y-axis measures the percent value-added to equityholders in this optimally-levered firm. The curves represent the value-added to equityholders that stems from 25% risk-reduction by increasing the optimal capital structure value (OCS), decreasing the maximal cost of the underinvestment problem (UI) or decreasing bankruptcy costs (BC). These are the solid, thickly-dashed and thinly-dashed lines, respectively.

Likewise, the effect of decreased volatility on the cost of making a larger amount of debt riskless is ambiguous. As we have already argued, the maximal cost of the underinvestment problem is exactly that—the cost to shareholders of making a fixed amount of risky debt risk-free. When the firm reduces its risk, that fixed amount of risky debt increases for a capital-structure-optimizing firm.

Figures (7) and (8) help illustrate the value to equityholders of the three leverage-based hedging motives. In figure (7), the firm optimally relevers as it reduces risk. In figure (8), the firm maintains a static capital structure. One might think that the underinvestment and bankruptcy cost stories would become much more valuable when the firm doesn't relever after hedging. This is not always the case. As the firm reduces risk and maintains a constant coupon outstanding, its endogenous default level rises. This is because its equity value declines, generally, as volatility is reduced, thus making the sale of equity less lucrative and incapable of financing coupon payments near the former, lower bankruptcy level. Therefore, when bankruptcy occurs, it will be more costly. In fact, note in the bottom-left figure of figure (8) that when volatility

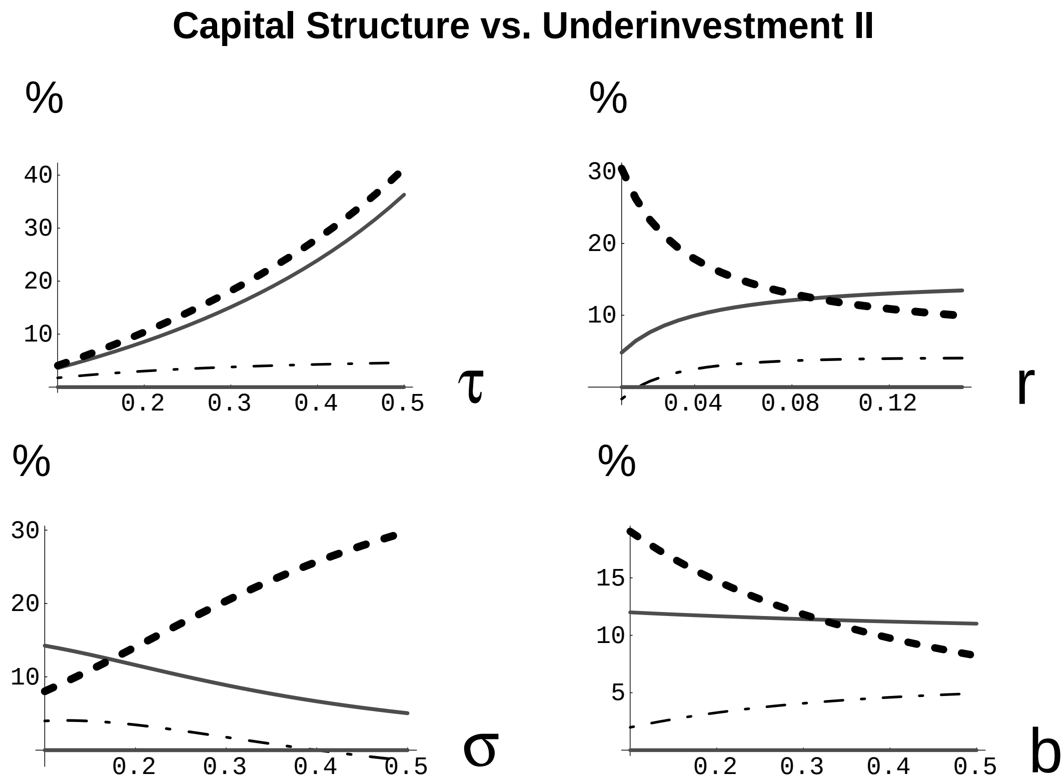


Figure 8: Everything is the same as in the previous figure except that the initially-optimally-levered firm no longer relevers after hedging for the purposes of measuring the value of risk-reduction to reducing bankruptcy costs and the maximal cost of the underinvestment problem.

Monotonicity in Risk

Story	Debt Re-Optimized	Constant Coupon
OCS	√	N/A
BC		
UI		

Table 8: This table illustrates which leverage motives of hedging are always beneficial. Abbreviations: OCS–optimal capital structure value-enhancement, BC–bankruptcy cost-reduction, UI–underinvestment cost-reduction.

is high, risk-reduction can increase bankruptcy costs for a given coupon outstanding. This conforms with Leland's result that junk-bonds can increase in value with risk and also leads us to corollary (4).

Corollary 4 *For a firm that maintains a constant amount of debt*[35] *neither the maximal cost of the underinvestment problem nor bankruptcy costs are monotonically de-*

[35] as measured by its coupon.

creasing in risk-reduction.

Proof: Using Leland (1994), it can be shown that:

$$\frac{\partial UI}{\partial \sigma_V} \; < \; 0 \, \forall \, \theta \text{ iff } (x \cdot g \, + \, 1) \cdot \log[h] \; > \; g \, \forall \, \theta$$

$$\frac{\partial BC}{\partial \sigma_V} \; < \; 0 \, \forall \, \theta \text{ iff } \log[\frac{1}{h}] \; > \; \frac{1}{x} \, \forall \, \theta$$

where $g \; = \; b[1 \Leftrightarrow r(1 \Leftrightarrow \tau)] \; + \; \tau(1 \Leftrightarrow b),$

$$h \; = \; \frac{Vr(1 \, + \, \frac{1}{x})}{C(1 \Leftrightarrow \tau)}$$

and θ is the vector of all relevant parameters.

Proposition (5) and corollary (4) are summarized in table (8).

5 A Review of Empirical Evidence

Since FASB rules have required off-balance sheet reporting of derivative positions, there has been a mushrooming of empirical work on the characteristics of firms that use derivatives. The significance of leverage as an explanatory factor in firms' derivative use runs the gammut from highly significant to not at all.

Economic methodology generally requires rejection of a normative theory when rational agents are observed to violate the theory's prescription. When principal-agent problems exist, what is normative for the principal may not be observed of the agent. Under these circumstances, empirical observation of the agent's behavior cannot compel rejection of a normative theory.

When a manager owns much of his employer's equity, he is averse to its idiosyncratic risk; this leads him to hedge and to underlever. When a manager owns stock options, he may be less averse to such idiosyncratic risk.[36] Employee stock options can thus mitigate the problem of underlevering but, if overused, can induce a manager to both lever and *speculate*, rather than to *hedge*, with derivatives. Since our normative theory calls for the use of derivatives for *hedging* to increase optimal leverage, the principal-agent problem may cause empirics to conflict with our prescription. This conflict between the positive and the normative, as explained, can occur no matter whether the agent is marginally more or less risk averse than the principal and is due to his non-diversified stake in the firm.

Despite this caveat, the evidence is largely favorable. Dolde (1995) finds that leverage is an insignificant explanatory variable for derivative use, but becomes significant when currency, commodity and interest rate risks are controlled for.[37] Among firms that speculate, however, leverage does not explain derivative use, further supporting

[36] Tufano (1995) shows that options significantly explain *negative* delta-hedging.

[37] See his page 201.

our theory.[38] This dichotomy between firms that speculate and firms that hedge is indirectly verified by Géczy, Minton and Schrand (1995). They report leverage to be the second most significant variable[39] in explaining interest rate derivative use (p-value $= 0.0004$) but only the ninth most significant in explaining foreign exchange derivative use (p-value $= 0.04$).[40] Perhaps the lower significance of leverage in explaining foreign exchange derivative use is due to their observation that speculators are more likely than hedgers to use currency derivatives, while comparably likely to use interest rate and commodity derivatives.[41]

Tufano (1995) finds leverage to be extremely significant (p-value < 0.01)[42] in explaining delta-hedging by North American gold mining firms when he controls for heteroskedasticity.[43]

Hentschel and Kothari (1995) offer some of the the most indirect, yet applicable, support for our theory. They show equity's β and σ to be only slightly increasing in the decile of derivative use, while leverage increases dramatically in derivative-use decile. This is strongly consistent with our hedging story. As firms aspire to higher optimal leverage, they must hedge and so dampen the leverage's effect on the risk of their equity.[44] One might wonder whether equity volatility causes derivative use or if the opposite is true. Hentschel and Kothari (H&K) shed light on this, as well. They show that the intercept on equity volatility is virtually invariant to derivative use and that the intensity of derivative use alone is insignificantly positively related to equity volatility. When leverage is added to the regression, derivative holdings become much more significant in explaining equity volatility and, more enlightening, are negatively related to equity volatility. In sum, the riskier a firm's equity, the more it uses derivatives. When controlling for leverage, the less a firm uses derivatives, the riskier its equity.[45]

Figure (9) predicts H&K's finding. In the top figure, equity volatility is only weakly related to firm volatility for a firm that is optimally-levered at each value of firm volatility, which can be considered a proxy for lack of derivative use. If the underlying firm volatilities of the firms H&K examined were not too high, then they indeed should have found a weak positive relationship between equity volatility and derivative use.[46] In the bottom figure, we see how equity volatility varies with firm volatility for firms that don't adjust their coupon outstanding with changes in leverage. Note that when leverage is controlled for in this manner, we find that equity

[38] See his Table 5C.

[39] Out of 15 explanatory variables.

[40] See their Tables 10 & 11.

[41] See their page 3 and their Table 7.

[42] See his Table V, Panel A.

[43] Tufano's most significant finding is the high significance of managerial stock and option ownership in explaining derivative use.

[44] See their Table 5.

[45] See their Table 3.

[46] Observe that the slope of the top figure is monotonically negative for firm volatilities below .12 and weakly negative for the group of firms with volatilities below some greater number—possibly as high as .25.

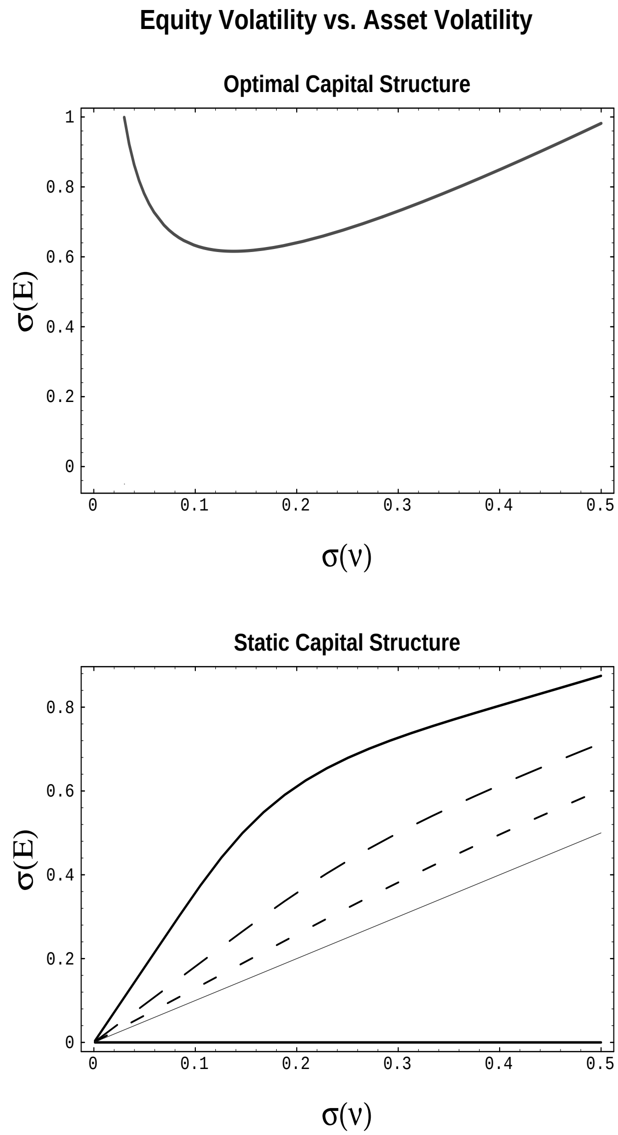


Figure 9: As usual, we have $V = \$100$, $r = 0.067$, $\tau = 0.25$ and $b = 0.22$. In the top figure, we show how equity volatility varies with firm (asset cum hedge portfolio) volatility for an optimally-levered firm. In the bottom figure, we show the same relationship for firms with debt of a constant coupon outstanding. For the solid, long-dashed, short-dashed and light-solid lines, respectively, coupons are \$6.44, \$4.00, \$2.00 and \$0. The \$6.44 coupon was chosen as the lowest optimal coupon for the range of volatilities examined. When firms are substantially more highly levered than is optimal, equity volatility is not well-behaved as a function of asset volatility.

volatility is strongly and positively related to firm volatility or strongly and negatively related to derivative use.

To summarize, when leverage is held fixed hedging will reduce equity volatility.

When firms hedge in order to increase leverage, hedging combined with optimal leverage can either reduce or increase equity volatility.

6 Conclusion

We have derived the optimal derivative portfolio and introduced the idea that firms' objective ought to be hedging the market value of their assets. Hedging individual risks is acceptable in our theory only as a means to that end. We have illustrated the robustness of our optimal derivative portfolio to variance-covariance misestimation and derived a number of interesting mathematical and econometric results.

We have demonstrated that hedging a firm's assets can result in an enhanced optimal capital structure, worth an extra 10% ⇔15% for current shareholders under very mild conditions. In the course of measuring the value of hedging to the capital structure, we have shown why leverage does indeed matter to shareholder value and we introduced a new methodology for estimating total direct and indirect bankruptcy costs.

We have offered a testable hypothesis that firms in competitive industries should be less likely to hedge their input costs than firms in oligopolistic industries. A series of testable hypotheses regarding conglomeration, spinoffs and the changing nature of mergers resulted from our theory of hedging motives. Our hedging story also has yielded a proposition on the dominance of callable debt over non-callable debt.

We have cited recent empirical evidence to support our story and to illustrate how principal-agent problems can intercede. Finally, we have contended that our story for the leverage motive of hedging offers more value to shareholders in most cases than competing leverage stories.

Appendix A

Proof of Proposition 1

Notation:

- TB, tax benefits
- BC, bankruptcy costs
- D, the price of debt at issuance
- P, the promised net present value of the debt. This amount exceeds the face value, a fact that compensates bondholders for the possibility that the firm will be unable to meet its obligation
- v, the mean preserving spread about the firm's mean value, 1
- b, the proportional bankruptcy cost suffered upon default, as a percentage of firm value
- η, the percentage of debt's value deductible to the corporation
- p_s, the state-price of state s
- NPV_s, the firm's state-s-contingent net-present value
- B, the set of all states of bankruptcy, when the firm is unable to pay bondholders promised amount P
- $f(s)$, the risk-neutral probability density function.

$$TB \Leftrightarrow BC = \tau\eta D \Leftrightarrow (1 \Leftrightarrow \tau) b \sum_{s \in B} p_s NPV_s$$

$$D = P \Leftrightarrow b \sum_{s \in B} p_s (P \Leftrightarrow (1 \Leftrightarrow b) NPV_s)$$

$$TB \Leftrightarrow BC = \tau\eta P \left(1 \Leftrightarrow \sum_{s \in B} p_s\right) + (\tau\eta \Leftrightarrow b) \sum_{s \in B} p_s NPV_s$$

Assume that, under the risk-neutral probability measure, Q, $NPV_s \sim U(1 \Leftrightarrow v, 1+v)$, and assume a riskless rate of zero.

$$\max_P \; TB \Leftrightarrow BC \Leftrightarrow \tag{A.1}$$

$$\max_P \tau\eta P \left[1 \Leftrightarrow \int_{1-v}^{P} f(s)ds\right] + (\tau\eta \Leftrightarrow b) \int_{1-v}^{P} f(s)NPV(s)ds \tag{A.2}$$

$$\text{Note that: } f(s) = \frac{1}{2v} \tag{A.3}$$

$$\frac{\partial(TB \Leftrightarrow BC)}{\partial P} = 0 \tag{A.4}$$

$$\Rightarrow P^* = \frac{1+v}{(1 + b)} \tag{A.5}$$

Now, combining (A.2), (A.3) and (A.5), we have:

$$(TB \Leftrightarrow BC)^* = \tag{A.6}$$

$$\tau\eta \frac{1+v}{(1 + b)} \left[1 \Leftrightarrow \frac{\frac{1+v}{(1+b)} \Leftrightarrow (1 \Leftrightarrow v)}{2v}\right] \tag{A.7}$$

$$+ \frac{\tau\eta \Leftrightarrow b}{4v} \left[\frac{3(1+v)^2}{(1 + b)^2} + (1 \Leftrightarrow v)^2 \Leftrightarrow \frac{4(1+v)(1 \Leftrightarrow v)}{(1 + b)}\right] \tag{A.8}$$

In order for the value of optimal capital structure to be monotonically decreasing in asset volatility, we need:

$$\begin{aligned} &\frac{\partial(TB \Leftrightarrow BC)^*}{\partial v} = \\ &2\tau\eta v \frac{b}{(1 + b)^2} + \frac{\tau\eta(1 \Leftrightarrow b)}{(1 + b)^2} + \frac{\tau\eta \Leftrightarrow b}{4} \left[\frac{3}{(1 + b)^2} \Leftrightarrow \frac{4}{(1 + b)} + 1\right] (1 \Leftrightarrow \frac{1}{v^2}) \\ &\leq 0 \; \forall \; v, \; \tau, \; \eta \,\&\, b \in (0,1). \end{aligned} \tag{A.9}$$

But necessary condition (A.9) is not always met.

Appendix B

Derivation of the Optimal Hedge Portfolio

$$\min_{\mathcal{D}} \mathcal{D}'\Omega\mathcal{D} \text{ s.t. } V = R\mathcal{D} \text{ where } R = [1 \mid \underline{0}']$$

$$\mathcal{L} = \mathcal{D}'\Omega\mathcal{D} + 2\gamma(R\mathcal{D} \Leftrightarrow V)$$

$$\begin{aligned}
&\text{First Order Conditions:} \\
\frac{\partial \mathcal{L}}{\partial \mathcal{D}^*} &= 2\Omega\mathcal{D} + 2R'\gamma = 0 \\
\frac{\partial \mathcal{L}}{\partial \gamma} &= 2(R\mathcal{D} \Leftrightarrow V) = 0 \\
&\text{Second Order Condition:} \\
\frac{\partial^2 \mathcal{L}}{\partial (\mathcal{D}^*)^2} &= \Omega, \\
\Omega \text{ is a VC } \Rightarrow |\Omega| > 0 &\Rightarrow \text{ minimization.}
\end{aligned}$$

Define:

$$W \equiv \begin{bmatrix} \Omega & R' \\ R & 0 \end{bmatrix},\ d \equiv \begin{bmatrix} \mathcal{D}^* \\ \gamma^* \end{bmatrix},\ v \equiv \begin{bmatrix} \underline{0} \\ V \end{bmatrix}$$

FOCs $\Rightarrow Wd = v \Rightarrow d = W^{-1}v$. Thus,

$$\begin{bmatrix} \mathcal{D}^* \\ \gamma^* \end{bmatrix} = \begin{bmatrix} \Omega & R' \\ R & 0 \end{bmatrix}^{-1} \begin{bmatrix} \underline{0} \\ V \end{bmatrix}$$

$$W = \begin{bmatrix}
\sigma_1^2 & . & . & . & \sigma_{n+1,1}\Delta_{n+1} & 1 \\
. & . & . & . & . & 0 \\
. & . & . & . & . & 0 \\
. & . & . & . & . & 0 \\
\sigma_{1,n+1}\Delta_n & . & . & . & \sigma_{n,n+1}\Delta_n\Delta_{n+1} & 0 \\
\sigma_{1,n}\Delta_n & . & . & . & \sigma_{n+1}^2\Delta_{n+1}^2 & 0 \\
1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

$$\begin{aligned}
W^{-1} &= \frac{C^T(W)}{|W|} = \frac{C(W)}{|W|} \\
|W| &= 1 \cdot (\Leftrightarrow 1)^{(n+2)+1} \cdot |W_{-1,n+2}| \\
|W_{-1,n+2}| &= 1 \cdot (\Leftrightarrow 1)^{(n+1)+1} \cdot |W_{-1\&n+2,1\&n+2}| \\
W_{-1\&n+2,1\&n+2} &\equiv \Omega_{-1,1} \\
|W| &= \Leftrightarrow |\Omega_{-1,1}|
\end{aligned}$$

Now that we have solved for $|W|$, we must solve for $C(W)$ to complete our calculation of W^{-1}. Note, however that $\mathcal{D}^*$ is simply the vector comprised by the first $n+1$ elements of the last column of W^{-1} times firm asset value, V. Therefore, we need only calculate the first $n+1$ elements of the last column of $C(W)$. Dividing that column by $|W|$ and multiplying it by V will give us $\mathcal{D}^*$. Suppressing the argument of $C(W)$, we have:

$$
\begin{aligned}
C_{n+2,j} &= (\Leftrightarrow 1)^{(n+2)+j} \cdot |W_{-(n+2),j}| \\
|W_{-(n+2),j}| &= 1 \cdot (\Leftrightarrow 1)^{(n+1)+1} \cdot |W_{-1\&(n+2),j\&(n+2)}| \\
W_{-1\&(n+2),j\&(n+2)} &\equiv \Omega_{-1,j} \\
C_{n+2,j} &= (\Leftrightarrow 1)^{j} |\Omega_{-1,j}| \\
\Rightarrow \mathcal{D}_j^* &= V \cdot \frac{(\Leftrightarrow 1)^{j} |\Omega_{-1,j}|}{\Leftrightarrow |\Omega_{-1,1}|} = V \cdot \frac{(\Leftrightarrow 1)^{j+1} |\Omega_{-1,j}|}{|\Omega_{-1,1}|} \\
&= V \cdot \frac{\lambda}{|\Omega_{-1,1}|}
\end{aligned}
$$

Appendix C

Derivation of the Maximal Risk-Reduction Achievable

From Appendix B, we know that

$$\sigma^* \;=\; \frac{1}{V}\sqrt{\mathcal{D}^{*\prime}\,\Omega\,\mathcal{D}^*} \;=\; \frac{\sqrt{\lambda'\,\Omega\,\lambda}}{|\Omega_{-1,1}|}$$

Lemma 1 *Consider the set of invertible matrices A and $C(A)$. $A'_i C^T_j \;=\; 0\;\forall\, i \neq j$.*

Proof:

$$\begin{aligned} A^{-1} &= \frac{C^T}{|A|} \\ AA^{-1} &= AC^T\frac{1}{|A|} \;=\; I \\ AC^T &= |A|I \end{aligned}$$

Note that $|A|I$ has $|A|$ on the diagonal and 0 elsewhere. The diagonal is generated by $A'_{i=j}C^T_j$ and the off-diagonals, which equal zero, are generated by $A'_{i\neq j}C^T_j$.

Now we shall continue with our calculation of the maximum achievable risk-reduction.

$$\lambda'\,\Omega \;=\; [|\Omega| \mid \underline{0}']$$

by lemma (1), since λ is the first row of $C(\Omega)$.

$$
\begin{aligned}
\lambda'\Omega\lambda &= |\Omega| \cdot \lambda_1 \;=\; |\Omega| \cdot |\Omega_{-1,1}| \\
\sigma^* V &= \frac{\sqrt{\lambda'\,\Omega\,\lambda}}{|\Omega_{-1,1}|} \;=\; V\,\frac{\sqrt{|\Omega| \cdot |\Omega_{-1,1}|}}{|\Omega_{-1,1}|} \\
&= V\sqrt{\frac{|\Omega|}{|\Omega_{-1,1}|}} \\
\text{Noting that: } |\Omega| &= \left(\prod_{i=1}^{n+1} \Delta_i^2 \sigma_i^2\right) |\rho| \\
\text{and } |\Omega_{-1,1}| &= \left(\prod_{i=2}^{n+1} \Delta_i^2 \sigma_i^2\right) |\rho_{-1,1}| \\
\Rightarrow \frac{|\Omega|}{|\Omega_{-1,1}|} &= \frac{|\rho|}{|\rho_{-1,1}|} \cdot \Delta_1^2 \sigma_1^2 \;=\; \frac{|\rho|}{|\rho_{-1,1}|} \cdot \sigma_1^2, \\
\sigma^* V &= V\sqrt{\frac{|\rho|}{|\rho_{-1,1}|} \cdot \sigma_1^2} \\
\sigma^* &= \sigma_1 \sqrt{\frac{|\rho|}{|\rho_{-1,1}|}} \equiv \sigma_1(1 \Leftrightarrow z)
\end{aligned}
$$

Appendix D

Interesting Results about Correlation Matrices and OLS

Theorem 1 *Consider correlation matrix ρ of dimension $n+1$. Consider, as well, matrix $\rho_{-i,i}$, which is the original matrix ρ with row and column i or sets of rows and columns i removed. The determinant of the reduced matrix is at least as large as that of the original.*

Proof: Note that, after performing our risk minimization routine, the risk of the firm is multiplied by the square root of a factor $\frac{|\rho|}{|\rho_{-1,1}|} \leq 1$, by the second order condition. Since the underlying asset of any of the hedge instruments over whose weights we are optimizing may be considered to be the firm, any identical row and column vector of the matrix ρ may be deleted in creating the matrix in the denominator. Like order statistics of the deleted row and column vector guarantee that they're identical, by the symmetry of $|\rho|$. Since we can repeat this deleting process up to n times, with the determinant of the remaining matrix never decreasing, i may be considered to be a set of rows and columns. In summary, we have:

$$
\begin{aligned}
&|\rho| \leq |\rho_{-i,i}| \\
&\Rightarrow |\Omega| \leq \left(\prod_i \Delta_i^2 \sigma_i^2\right) |\Omega_{-i,i}|
\end{aligned}
$$

This is equivalent to the statement that $\prod_{j \neq i} C_{i,j}(\rho) \cdot |\rho_{-i,j}| \leq 0$ since $C_{i,i}(\rho) \cdot |\rho_{-i,i}| = |\rho_{-i,i}|$ and $\prod_j C_{i,j}(\rho) \cdot |\rho_{-i,j}| = |\rho|$.

Theorem 2 *Consider the ratio defined in theorem (1). Now consider reducing the original matrix $|\rho|$ by eliminating rows and columns j. Take the ratio of the reduced matrix to its further reduced counterpart (i.e. sans rows and columns $i \neq j$). The latter ratio is at least as large as the former.*

Proof: Since we can only reduce risk further by hedging over previously available assets plus extra assets j. This risk-reduction will be non-zero, provided the weights on the new assets are not all zero i.e. if $\sum_j |\Omega_{-1,j}|^2 \neq 0$. Since the optimization can be performed over assets all having variance 1, the following inequality will be strict if and only if $\sum_j |\rho_{-1,j}|^2 \neq 0$.

$$
\begin{aligned}
&\sqrt{\frac{|\rho|}{|\rho_{-i,i}|}} \leq \sqrt{\frac{|\rho_{-j,j}|}{|\rho_{-i\&j,i\&j}|}} \\
&\Leftrightarrow |\rho| \cdot |\rho_{-i\&j,i\&j}| \leq |\rho_{-i,i}| \cdot |\rho_{-j,j}|
\end{aligned}
$$

Define the first of $n = 1$ variables to be the dependent variable of a regression and define the other variables as explanatory.

Definition 1 *A* centered *OLS is one in which the means of all variables are zero, i.e. the mean of each explanatory variable's observation vector is subtracted from each element in the vector prior to regressing; likewise with the independent variable's observation vector.*

Theorem 3 *The "true" β vector of a centered multivariate OLS can be described as follows:*

$$\beta_{-1} = \Leftrightarrow \frac{C_{1,-1}(\Omega)}{|\Omega_{-1,1}|} \cdot V = \Leftrightarrow \Omega^{-1}_{1,-1} \cdot \frac{\Omega}{|\Omega_{-1,1}|} \cdot V = \Leftrightarrow \mathcal{D}^{*}.$$

And for any individual regression coefficient, [47]

$$\beta_j = (\Leftrightarrow 1)^j \cdot \frac{|\rho_{-1,j}|}{|\rho_{-1,1}|} \cdot \frac{\sigma_1 V}{\sigma_j} \; \forall \, j > 1.$$

Proof: Note from proposition (6) that when Ω is known, the optimal hedge instrument weights are analogous to the negative of the coefficients of a centered OLS regression. Note that when Ω is known, the true β is also known, as a centered OLS captures only second moment effects.

Given proposition (6), we can proceed to prove our claimed property of β_j by discussion of $\mathcal{D}^*_j$.

$$\begin{aligned}
\mathcal{D}^* &= \frac{V}{|\Omega_{-1,1}|} \cdot \lambda \\
\mathcal{D}^*_j &= \frac{V}{|\Omega_{-1,1}|} \cdot \left((\Leftrightarrow 1)^{j+1} \cdot |\Omega_{-1,j}|\right) \\
&= \frac{V}{\left(\prod_{j=2}^{n+1} \sigma_j^2\right) \, |\rho_{-1,1}|} \cdot \frac{\left((\Leftrightarrow 1)^{j+1} \cdot \prod_{j=1}^{n+1} \sigma_j^2\right) \, |\rho_{-1,j}|}{\sigma_1 \sigma_j} \\
&= (\Leftrightarrow 1)^{j+1} \cdot \frac{|\rho_{-1,j}|}{|\rho_{-1,1}|} \cdot \frac{\sigma_1 V}{\sigma_j} = \Leftrightarrow \beta_j \;\; \forall \; j > 1.
\end{aligned}$$

Theorem 4 *The R-squared of a centered regression converges to one minus the ratio of the determinant of the correlation matrix of the independent* and *explanatory variables to the determinant of the correlation matrix of explanatory variables.*

Proof: In proposition (6), we performed the minimization over β for the expectation of the squared error of a single observation. Here, we do so for N-observations.

[47] Note that, in a conventional centered multivariate OLS, we are only concerned with coefficients on the explanatory variables, which are elements $j = 2, ..., n+1$ of the vector $-\frac{C_1(\Omega)}{|\Omega_{-1,1}|} \cdot V$.

$$
\begin{aligned}
(\sigma^* V)^2 \cdot N &= \min_{\beta} \left[(y \Leftrightarrow X\beta)'(y \Leftrightarrow X\beta) \right] \\
\sigma^* V \sqrt{N} &= \sqrt{\varepsilon' \varepsilon} \;=\; \sqrt{\mathrm{SSE}} \\
\text{Note that SST} &= \mathrm{SSR} \;+\; \mathrm{SSE},\; R^2 \equiv 1 \Leftrightarrow \frac{\mathrm{SSE}}{\mathrm{SST}}, \\
s_y^2 &= \frac{\mathrm{SST}}{N \Leftrightarrow K} \text{ and } s_{\hat{\varepsilon}}^2 \;=\; \frac{\mathrm{SSE}}{N \Leftrightarrow K} \\
R^2 &= 1 \Leftrightarrow \frac{\mathrm{SSE}}{\mathrm{SST}} \\
&= 1 \Leftrightarrow \left(\frac{\mathrm{SSE}/(N \Leftrightarrow K)}{SST/(N \Leftrightarrow 1)} \right) \cdot \frac{N \Leftrightarrow K}{N \Leftrightarrow 1} \\
&= 1 \Leftrightarrow \frac{s_{\hat{\varepsilon}}^2}{s_y^2} \cdot \frac{N \Leftrightarrow K}{N \Leftrightarrow 1}.
\end{aligned}
$$

By a generalization of the Slutzky Theorem,

$$
\begin{aligned}
\operatorname{plim} \frac{s_{\hat{\varepsilon}}^2}{s_y^2} \cdot \frac{N \Leftrightarrow K}{N \Leftrightarrow 1} &= \frac{\sigma_{\varepsilon}^2}{\sigma_y^2} \equiv \frac{(\sigma^* V)^2 \cdot N}{(\sigma_1 V)^2 \cdot N} \;=\; \left(\frac{\sigma^*}{\sigma_1} \right)^2 \\
\Rightarrow \operatorname{plim} R^2 &= 1 \Leftrightarrow \left(\frac{\sigma^*}{\sigma_1} \right)^2
\end{aligned}
$$

where:
SSE is the sum of squared errors,
SST is the total sum of squares,
SSR is the regression sum of squares,
K is the number of explanatory variables
s_y^2 is the unbiased estimator for the variance of the independent variable, y and
$s_{\hat{\varepsilon}}^2$ is the unbiased estimator for the variance of the estimated residuals.

Recalling that

$$
\begin{aligned}
\sigma^* \;=\; \sigma_y \sqrt{\frac{|\rho|}{|\rho_{-y,y}|}}, \\
\Rightarrow \left(\frac{\sigma^*}{\sigma_1} \right)^2 &= \frac{\sigma_1^2 \cdot \frac{|\rho|}{|\rho_{-1,1}|}}{\sigma_1^2} \;=\; \frac{|\rho|}{|\rho_{-1,1}|} \;=\; \frac{|\rho|}{|\rho_{-y,y}|} \\
\Rightarrow \operatorname{plim} R^2 &= 1 \Leftrightarrow \frac{|\rho|}{|\rho_{-y,y}|}.
\end{aligned}
$$

Equally, since $plim\ \hat{\beta} \;=\; \beta$,

$$
\operatorname{plim} \hat{\beta}_{-1} \;=\; \Leftrightarrow \frac{C_{1,-1}(\Omega)}{|\Omega_{-1,1}|} \cdot V \;=\; \Leftrightarrow \mathcal{D}^*_{-1}
$$

Proposition 6 *When Ω is known, the optimal hedge portfolio consists in taking a position of $\Leftrightarrow V\beta$ in each of the hedge instruments, where β are analogous to weights from a multivariate least squares regression of the firm's assets' returns on hedge instrument returns.*

Proof: Simply restate the optimization problem as follows:

$$\min_{\beta} \mathrm{E}[(y \cdot V \;+\; X'(\Leftrightarrow\beta_{-1}))'(y \cdot V \;+\; X'(\Leftrightarrow\beta_{-1}))]$$

where y is the deviation of the firm's return from its expectation , X is the vector of the deviation of the n hedge instrument returns from their means and $\Leftrightarrow\beta_{-1}$ is the vector of hedge instrument portfolio weights, in dollars of notional underlying.

This is simply an ordinary multivariate least squares in which the sign of the factor loadings is opposite the convention. Consider that second moments constitute sufficient statistics for an OLS optimization (i.e OLS cannot yield more accurate coefficient estimates than when the observations reveal the true distribution of the explanatory variables). Consider, as well, that omitting the firm as an explanatory variable is the same as constraining our exposure to the firm's underlying assets to the unhedged exposure. Then we have minimized the variance of the firm's hedged returns in the same manner as we did in the original optimization problem.

More rigorously, we have:

$$\min_{\beta} \mathrm{E}[(y \cdot V \;\Leftrightarrow X'\beta_{-1})'(y \cdot V \;\Leftrightarrow X'\beta_{-1})]$$

$$\min_{\beta} \mathrm{E}[((y \mid X')(\Leftrightarrow\beta))'(y \mid X')(\Leftrightarrow\beta)] \text{ s.t. } \Leftrightarrow\beta_1 \;=\; V$$

Now, let $\mathcal{D} \;=\; \Leftrightarrow\beta$ and define $H_{-1} \;=\; X$.

$$\min_{\mathcal{D}} \mathrm{E}[(y \cdot V \;+\; (H'_{-1}\mathcal{D}_{-1}))'(y \cdot V \;+\; (H'_{-1}\mathcal{D}_{-1}))]$$

Let $H_1 \equiv y$ and perform the following minimizations subject to the constraint that $\mathcal{D}_1 \;=\; V$.

$$\begin{aligned}
& \min_{\mathcal{D}} \mathrm{E}[(H'\mathcal{D})'(H'\mathcal{D})] \Leftrightarrow \min_{\mathcal{D}} \mathrm{E}[(\mathcal{D}'H)'(H'\mathcal{D})] \\
\Leftrightarrow \quad & \min_{\mathcal{D}} \mathrm{E}[(\mathcal{D}'H)(H'\mathcal{D})] \Leftrightarrow \min_{\mathcal{D}} \mathrm{E}[\mathcal{D}'HH'\mathcal{D}] \\
\Leftrightarrow \quad & \min_{\mathcal{D}} \mathcal{D}'\mathrm{E}[HH']\mathcal{D} \Leftrightarrow \min_{\mathcal{D}} \mathcal{D}'\Omega\mathcal{D}
\end{aligned}$$

Corollary 5 *When Ω is unknown, a multivariate least squares regression may be used to find portfolio weights only for hedge instruments with a constant volatility.*

Sketch of Proof: Since our goal is to find the optimal instantaneous hedge portfolio, we are concerned with the variance-covariance matrix at the current time. A multivariate least squares regression that includes assets with non-constant volatilities will recommend weights that are not based on the current volatility.[48] Such an approach would be more acceptable if the goal was to find the best static hedge for a forthcoming time period comparable to that of the data's timespan. But certainly, such an approach would want for accuracy.

It would be far better, if implementing a dynamic hedge, to estimate local variances, correlations and deltas using implied volatilities, history and pricing models, respectively. Latané and Rendleman (1976) has shown that Black Scholes estimates of implied volatility do a better job than historical volatility in forecasting future volatility.[49] With regard to estimating correlations, recent work by Bodurtha and Shen (1994) and by Marsh and Mayhew (1996) on estimating implied correlations can further aid in estimating Ω. If the hedge instrument in question is a future, options on that future may be used to obtain implied volatilities. If the hedge instrument in question is an option, option pricing models can yield the implied volatility and delta. If the hedge instrument is a bond future, option or swap, then fixed income pricing models may be applied. In estimating the correlation between the firm's assets and hedge instruments, the return on the firm's assets will not be directly observable if the firm is levered. The correlation between the levered firm and a hedge instrument is the same as that between the hedge instrument and the firm's assets. Being able to directly observe the firm's asset returns would provide for more accurate correlation estimates. Such asset returns may be backed out of levered equity returns by applying Longstaff and Schwartz (1995) or Leland and Toft (1996). These models may also be used to calculate implied volatility on a levered firm's underlying assets even when no options are traded.

Corollary 6 *The optimal weight on hedge instrument j is independent of all $\sigma_{i \neq j, i \neq 1}$, is directly proportional to σ_1 and is inversely related to σ_j.*

Proof: By proposition (6), hedge instrument weights are analogous to $\Leftrightarrow \beta_{OLS}$. By theorem (3) , β_j of a centered OLS is directly proportional to σ_1, is inversely related to σ_j and is independent of all other hedge instrument volatilities.

[48]Even if the underlying assets of hedge instruments have constant volatility, the hedge instruments themselves may not. Options, for example, have varying deltas, in which their volatility is locally linear for many models. When volatility does vary, a generalized least squares may be employed, using intraday volatility data, though even this modification to the regression does not make the regression approach attractive.

[49]There is some controversy over this finding. See Figlewski (?) and Geske (?).

References

Altman, Edward I. and Joseph C. Bencivenga, 1995. "A Yield Premium Model For T he High-Yield Debt Market." *Financial Analysts Journal*, September-October: 49-56.

Berger, Philip G. and Eli Ofek, 1995. "Diversification's Effect on Firm Value. " *Journal of Financial Economics*, 37: 39-65.

Berger, Philip G. and Eli Ofek, 1995. "Bustup Takeover of Value-Destroying Diversified Firms." NYU Working Paper: June revision.

Block, Stanley B. and Timothy J. Gallagher, 1986. "The Use of Interest Rate Fu tures and Options by Corporate Financial Managers." *Financial Management*, 15: 7 3-78.

Bodurtha, James N., Jr. and Qi Shen, 1994. "Implied Covariance in PHLX Deutschemark and Yen OPtion Values." University of Michigan Working Paper: Nov ember Revision.

Constantinides, George M. and Jonathan E. Ingersoll, Jr., 1984. "Optimal Bond Trading with Personal Taxes." *Journal of Financial Economics*, 13(3): 299-335.

DeMarzo, Peter M. and Darrell Duffie, 1995 (1992). "Corporate Incentives for H edging and Hedge Accounting." Stanford University Working Paper, 1995 and 1992 version s.

Dolde, Walter, 1993. "Use of Foreign Exchange and Interest Rate Risk Managemen t in Large Firms." University of Connecticut Working Paper.

Dolde, Walter, 1995. "Hedging, Leverage, and Primitive Risk." *Journal of Financial Engineering*, 4(2): 187-216.

Froot, Kenneth A., David S. Scharfstein and Jeremy C. Stein, 1993. "Risk Management: Coordinating Corporate Investment and Financing Policies." *Journal of Finance*, 48: 1629-58.

Gardner, Grant W. and Douglas Stone, 1995. "Estimating Currency Hedge Ratios for International Portfolios." *Financial Analysts Journal*, November/December: 58-64.

Géczy, Christopher, Bernadette A. Minton and Catherine Schrand, 1995. "Why Firms Use Derivatives: Distinguishing Among Existing Theories." Ohio State University Working Paper, September Revision.

Green, William H., 1993. *Econometric Analysis*, Macmillan Publishing Compa ny. New York, NY.

Grossman, Sanford J. and Oliver D. Hart, 1980. "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation." *Bell Journal of Economics*, 11(1): 42-64.

Hart, Ian and Michael Ross, 1994. "Striking Continuity." *Risk Magazine*, 7(6): 51-54

Hentschel, Ledger and S.P.Kothari, 1995. "Are Corporations Managing or Taking Risks with Derivatives?." University of Rochester Working Paper, August Revision.

Jacquie, Laurent L., 1981. "Management of Foreign Exchange Risk: A Review Arti cle." *Journal of International Business Studies*, Spring/Summer:81-101.

Kaplan, Steven N and Michael S. Weisbach, 1992. "The Success of Acquirisitions : Evidence from Divestitures." *Journal of Finance*, 47(1): 107-138.

Lang, Larry H. P. and René M. Stulz, 1994. "Tobin's q, Corporate Diversification and Firm Performance." *Journal of Political Economy*, 102(6): 1248-1280.

Latané, Hery A. and Richard J. Rendleman, Jr., 1976. "Standard Deviations of Stock Price Ratios Implied in Option Prices." *Journal of Finance*, 31(2): 369-81.

Leland, Hayne E., 1989. "LBOs and Taxes: No One to Blame But Ourselves?." *California Management Review*, 32(1): 19-29.

Leland, Hayne E., 1994. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." *Journal of Finance*, 49(4): 1213-1252.

Leland, Hayne E. and Klaus Bjerre Toft, 1996. "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads." *forthcoming Journal of Finance*, 51(3).

Litzenberger, Robert H. and Krishna Ramaswamy, 1979. "The Effect of Personal Taxes and Dividends on Capital Asset Prices." *Journal of Financial Economics*, 7(2): 163-95.

Longstaff, Francis A. and Eduardo S. Schwartz, 1995. "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt." *Journal of Finance*, 50(3): 789-819.

Marsh, Terry and Stewart Mayhew, 1996. University of California at Berkeley. W ork in Progress.

Mayers, David and Clifford W. Smith, Jr., 1982. "On the Corporate Demand for Insurance: Evidence from the Reinsurance Market." *Journal of Business*, 55(2): 281-96.

Mayers, David and Clifford W. Smith, Jr., 1987. "Corporate Insurance and the U nderinvestment Problem." *Journal of Risk and Insurance*, 54: 45-54.

Mayers, David and Clifford W. Smith, Jr., 1990. "On the Corporate Demand for Insurance: Evidence from the Reinsurance Market." *Journal of Business*, 63(1): 19-41.

Myers, Stewart, 1977. "Determinants of Corporate Borrowing." *Journal of Financial Economics*, 5(2): 147-75.

Nance, Deana R., Clifford W. Smith, Jr. and Charles W. Smithson, 1993. "On the Determinants of Corporate Hedging." *Journal of Finance*, 48(1): 267-284.

Servaes, Henri, 1995. "The Value of Diversification during the Conglomerate Me rger Wave." Forthcoming *Journal of Finance.*

Smith, Clifford W., Jr., 1995. "Corporate Risk Management: Theory and Practice ." *Journal of Derivatives*, Summer: 21-30.

Smith, Clifford W. and René M. Stulz, 1985. "The Determinants of Firm's Hedging Policies." *Journal of Financial and Quantitative Analysis*, 20(2):391-405.

Stulz, René M., 1984. "Optimal Hedging Policies." *Journal of Financial and Quantitative Analysis*, 19(2):127-140.

Tufano, Peter, 1995. "Who Manages Risk? An Empirical Examination of Risk Management Practices in the Gold Mining Industry." Harvard University Working Paper, December Revision.

Wall, Larry D. and John Pringle, 1989. "Alternative Explanations of Interest Rate Swaps: An Empirical Analysis." *Financial Management*, 18: 59-73.

Warner, Jerold B., 1977. "Bankruptcy Costs: Some Evidence." *Journal of Finance*, 32(2): 71-81.